

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2003, 01:30:31 : Search time 85 seconds  
(without alignments)  
805.774 Million cell updates/sec

Title: US-09-846-637C-4  
Perfect score: 2619  
Sequence: 1 MADYLISGCTSYVPPDGLTA.....SSAQVEGVFSLHSYEKRLIF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq.101002.\*

1:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2605	99.5	514	11	Human IMPDH, Hom
2	2605	99.5	514	23	Human wild-type, t
3	2595	99.1	514	23	Human wild-type, t
4	2595	99.1	514	23	Reference sequence
5	2595	97.7	514	11	Human hamster IM
6	2234	85.3	514	23	Human wild-type in
7	2229	85.1	514	23	Human type I inos
8	2227	85.0	514	23	Human type I IMPDH
9	2222	84.8	514	23	Human wild-type, t
10	2202	84.1	514	23	Human wild-type, t

11	2201	84.0	514	23	AAE18191	Human wild-type, t
12	2141.5	81.8	604	22	ABG22269	Novel human diapo
13	2115.5	80.8	509	23	AAU99361	Human oxidoreducta
14	1862	71.1	384	23	AAE18179	Human type II inos
15	1861.5	71.1	385	23	AAE18179	Human modified ino
16	1860.5	71.0	385	23	AAE18177	Human modified ino
17	1860	71.0	384	23	AAE18174	Human type II inos
18	1859.5	71.0	385	23	AAE18181	Human type II IMPD
19	1858	70.5	384	23	AAE18169	Human modified ino
20	1857.5	70.5	385	23	AAE18182	Human modified ino
21	1857.5	70.5	385	23	AAE18183	Human type II IMPD
22	1857	70.5	384	23	AAE18167	Human type II inos
23	1856	70.5	384	23	AAE18165	Human modified ino
24	1856	70.5	384	23	AAE18170	Human modified ino
25	1856	70.5	384	23	AAE18172	Human type II inos
26	1855.5	70.6	385	23	AAE18176	Human modified ino
27	1855.5	70.6	385	23	AAE18178	Human modified ino
28	1855.5	70.6	385	23	AAE18180	Human modified ino
29	1855	70.6	384	23	AAE18168	Human modified ino
30	1853	70.6	384	23	AAE18173	Human modified ino
31	1852	70.7	384	23	AAE18171	Human modified ino
32	1851.5	70.7	385	23	AAE18184	Human modified ino
33	1764.5	67.4	537	22	ABB58547	Drosophila melanog
34	1764.5	67.4	537	22	ABB65768	Drosophila melanog
35	1644	62.8	384	23	AAE18175	Human type I inos
36	1524	58.2	430	22	ABG23731	Novel human diapo
37	1210.5	46.2	371	20	AAV08965	A. gossypii inosin
38	1155	44.1	502	21	AAV30888	Arabidopsis thalia
39	1114.5	42.6	503	21	AAV20989	Arabidopsis thalia
40	1106.5	42.2	503	21	AAV43108	Arabidopsis thalia
41	1066	40.7	287	22	ABG07490	Novel human diapo
42	1043.5	39.8	403	22	AAV07079	S cerevisiae apopt
43	1043.5	39.0	443	21	AAV30889	Arabidopsis thalia
44	996	38.0	435	21	AAV30890	Arabidopsis thalia
45	987.5	37.7	444	21	AAV20990	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
AA05432  
ID AA05432 standard; protein; 514 AA.  
XX  
AC AA05432:  
XX  
DT 31-AUG-1990 (first entry)  
XX  
DE Human IMPDH.  
XX  
KW Inosine 5'-mono-phosphate dehydrogenase; hepatomas;  
KW guanosine monophosphate.  
XX  
OS Homo sapiens.  
XX  
PN WO9001545-A.  
XX  
PD 22-FEB-1990.  
XX  
PF 02-AUG-1989: 89WO-0000344.  
XX  
PR 12-AUG-1988: 88US-0232302.  
XX  
PA (ARCH-) ARCH. DEV. CORP.  
XX  
PI Collart FR, Huberman E;  
XX  
DR WPI: 1990-083504/11.  
XX  
DR P-PSDB; AA05432.  
XX  
PT DNA encoding eukaryotic inosine 5'-mono-phosphate dehydrogenase -  
PT esp. hepatomas, and to produce guanosine monophosphate.



RESULT 3  
AAE18186  
ID AAE18186 standard: Protein; 514 AA.  
XX AAE18186;  
AC AAE18186;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
XX  
DE Human wild-type, type II IMPDH #1.  
XX  
XX  
KM Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;  
KW proliferative-type disease; cancer; imaging methodology; cytostatic;  
KM enzyme; therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 111..243  
FT /label= sub\_domain  
XX  
XX WO200185952-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 10-MAY-2001: 2001WO-US15457.  
XX  
XX 10-MAY-2000: 2000US-203448P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff JE, Einspahr HM, Kish K;  
XX  
XX WPI: 2002-164105/21.  
XX  
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
XX  
PS Claim 55; Fig 2; 161pp; English.  
XX  
XX The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human wild-type type II IMPDH.  
XX  
XX  
SQ Sequence 514 AA:  
Query Match 99.1%; Score 2595; DB 23; Length 514;  
Best Local Similarity 99.2%; Pred. No. 2, 7e-245;  
Matches 510: Conservative 0; Mismatches 4; Gaps 0;  
Insets 0; Gaps 0;  
QY 1 MADYLLSGSTVYPPDGLTAQQLFNCGDGLFYNDPLILPGYIDFTAD VDLTSALTKKIT 60  
DB 1 MADYLLSGSTVYPPDGLTAQQLFNCGDGLFYNDPLILPGYIDFTAD VDLTSALTKKIT 60  
QY 61 LKTPPLVSSPMDDTYTEAGMALIAMLATGCGITIHNCCTPEQANFVRKV KYEGGFTIDPVV 120  
DB 61 LKTPPLVSSPMDDTYTEAGMALIAMLATGCGITIHNCCTPEQANFVRKV KYEGGFTIDPVV 120

QY 121 LSPKDRVDYFEAKARRHFGCIPITDNGHMSRLVGISSRDTDFLKEEHDCELEIMT 180  
DB 121 LSPKDRVDYFEAKARRHFGCIPITDNGHMSRLVGISSRDTDFLKEEHDCELEIMT 180  
QY 181 KREDLVVAPRSITLKEANETLQSRKKGKLPVNEDELVAIARTDLKKNRDYPLASKDA 240  
DB 181 KREDLVVAPRSITLKEANETLQSRKKGKLPVNEDELVAIARTDLKKNRDYPLASKDA 240  
QY 241 KROLLCGAATCTHEDDKRYRLDLLAQAGVDVYVLDSSQNSIFQINMIKIKKKYPLQYI 300  
DB 241 KROLLCGAATCTHEDDKRYRLDLLAQAGVDVYVLDSSQNSIFQINMIKIKKKYPLQYI 300  
QY 301 GGNVYVTAQAQKNLIDAGVDALRVGMGSGSICIIQEVLAGRPGATVYKYVEYARRFGVP 360  
DB 301 GGNVYVTAQAQKNLIDAGVDALRVGMGSGSICIIQEVLAGRPGATVYKYVEYARRFGVP 360  
QY 361 VIADGGIIONVGHIAKALALAGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDAM 420  
DB 361 VIADGGIIONVGHIAKALALAGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDAM 420  
QY 421 DKHLSSQNRFFSEADKIKVAQVSGAVODKGSIHKFPVYLIAGIQHSCODIGAKSLTOYR 480  
DB 421 DKHLSSQNRFFSEADKIKVAQVSGAVODKGSIHKFPVYLIAGIQHSCODIGAKSLTOYR 480  
QY 481 AMMYSGELKFEKRTSSAOYEGVHSLHSTEKRLF 514  
DB 481 AMMYSGELKFEKRTSSAOYEGVHSLHSTEKRLF 514  
RESULT 4  
AAU10695  
ID AAU10695 standard: Protein; 514 AA.  
XX  
XX AAU10695;  
XX  
XX 25-FEB-2002 (first entry)  
XX  
XX Reference sequence for human IMPDH2 polypeptide.  
XX  
XX  
XX Human; single nucleotide polymorphism: SNP; IMPDH2; chromosome 3p21.2;  
XX IMP dehydrogenase 2; haplotyping; genotyping; cancer; cytostatic.  
XX  
XX Homo sapiens.  
XX  
XX WO200177363-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 11-APR-2001: 2001WO-US11851.  
XX  
XX 11-APR-2000: 2000US-196248P.  
XX  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
XX Chew A, Choi JT, Koshy B, Lee HH, Stephens JC;  
PI WPI: 2002-041297/05.  
DR N-PSDB: AAS18240.  
XX  
XX New isolated polynucleotide having polymorphic variant of IMP2  
PT dehydrogenase gene, useful for studying expression of the gene in vivo,  
PT and for testing efficacy of therapeutic agents for cancer in biological  
PT system -  
XX  
XX  
PS Disclosure; Fig 3; 70pp; English.  
XX  
XX The present invention relates to novel single nucleotide polymorphisms  
CC (SNPs) in the human IMP dehydrogenase 2 (IMPDH2) gene located on  
CC chromosome 3p21.2, and methods for haplotyping and/or genotyping the  
CC IMPDH2 gene in an individual. The methods of the invention make use of  
CC allele-specific oligonucleotides (ASOs) as probes and primers and/or  
CC primer-extension oligonucleotides for detecting the IMPDH2 gene  
CC polymorphisms. The polynucleotides and screened compounds are

CC useful for (developing) treatment of diseases associated with IMPDH2  
CC activity, such as cancer. The present sequence represents a reference  
CC sequence for the IMPDH2 polypeptide.

XX Sequence 514 AA:

Query Match 99.1%; Score 2595; DB 23; Length 514;  
Best Local Similarity 99.2%; Pred. No. 2.7e-245;  
Matches 510; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 MADYILISGTSYVPDDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTKKIT 60
    |||
DB 1 MADYILISGTSYVPDDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTKKIT 60
OY 61 LKTPLYSSPMDIVTEAGMAIAMAALTGGIGFIHNCTPEFOANEVRKVKYEGGFTTDPVY 120
    |||
DB 61 LKTPLYSSPMDIVTEAGMAIAMAALTGGIGFIHNCTPEFOANEVRKVKYEGGFTTDPVY 120
OY 121 LSPKDRVDFEAKARHGFSGIPITDTGRMGSRVGISSRIDFLKEEHDFLEIMT 180
    |||
DB 121 LSPKDRVDFEAKARHGFSGIPITDTGRMGSRVGISSRIDFLKEEHDFLEIMT 180
OY 181 KREDLVVAARSTILKEANELLQSRKKGKLPYVNEDELVAIARTDLKKNRDPPLASKDA 240
    |||
DB 181 KREDLVVAARSTILKEANELLQSRKKGKLPYVNEDELVAIARTDLKKNRDPPLASKDA 240
OY 241 KROLGGAALGTHEDDKYRLDLAAGVGVVLDSSQGSIFQIMNKYIKDKYPMLOYI 300
    |||
DB 241 KROLGGAALGTHEDDKYRLDLAAGVGVVLDSSQGSIFQIMNKYIKDKYPMLOYI 300
OY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITQEVLAGCRPOATAVYKVEYARRFVGP 360
    |||
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITQEVLAGCRPOATAVYKVEYARRFVGP 360
OY 361 VIADGSIQVNGHIAKALALGASTVMMGSLAATTEARGEYFSDGIRLKKYRGMSLDM 420
    |||
DB 361 VIADGSIQVNGHIAKALALGASTVMMGSLAATTEARGEYFSDGIRLKKYRGMSLDM 420
OY 421 DKHLSSQNRYSFSEADKIKVAQGVSGAVDPDKGSIHKFVPLYLAGIQHSCODIGAKSLTOVR 480
    |||
DB 421 DKHLSSQNRYSFSEADKIKVAQGVSGAVDPDKGSIHKFVPLYLAGIQHSCODIGAKSLTOVR 480
OY 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEYKRLF 514
    |||
DB 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEYKRLF 514

RESULT 5
AAR05431
ID AAR05431 standard; protein; 514 AA.
XX
AC AAR05431;
XX
DT 31-AUG-1990 (first entry)
XX
DE Chinese hamster IMPDH.
XX
KW Inosine 5'-mono-phosphate dehydrogenase; hepatomas;
KM guanosine monophosphate.
XX
OS Cricetus sp.
XX
PN WO9001545-A.
XX
PD 22-FEB-1990.
XX
PF 02-AUG-1989; 89WO-0000344.
XX
PR 12-AUG-1988; 88US-0232302.
XX
PA (ARCH-) ARCH. DEV. CORP.
XX
PI Collart FR, Huberman E;

```

XX WPI; 1990-083504/11.  
DR N-PSDB; AA003541.  
XX  
PT DNA encoding eukaryotic inosine 5'-mono-phosphate dehydrogenase -  
PT used to detect high levels of enzyme activity of tumour cells,  
PT esp. hepatomas, and to produce guanosine monophosphate.  
PS Claim 8; Fig 2; 51pp; English.

CC A 35 aa segment corresponds to deduced aa residues 336-370 in  
CC both the human and Chinese hamster proteins.  
CC See also AA003540.

XX Sequence 514 AA:

Query Match 97.7%; Score 2559; DB 11; Length 514;  
Best Local Similarity 98.1%; Pred. No. 9.3e-242;  
Matches 504; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

OY 1 MADYILISGTSYVPDDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTKKIT 60
    |||
DB 1 MADYILISGTSYVPDDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTKKIT 60
OY 61 LKTPLYSSPMDIVTEAGMAIAMAALTGGIGFIHNCTPEFOANEVRKVKYEGGFTTDPVY 120
    |||
DB 61 LKTPLYSSPMDIVTEAGMAIAMAALTGGIGFIHNCTPEFOANEVRKVKYEGGFTTDPVY 120
OY 121 LSPKDRVDFEAKARHGFSGIPITDTGRMGSRVGISSRIDFLKEEHDFLEIMT 180
    |||
DB 121 LSPKDRVDFEAKARHGFSGIPITDTGRMGSRVGISSRIDFLKEEHDFLEIMT 180
OY 181 KREDLVVAARSTILKEANELLQSRKKGKLPYVNEDELVAIARTDLKKNRDPPLASKDA 240
    |||
DB 181 KREDLVVAARSTILKEANELLQSRKKGKLPYVNEDELVAIARTDLKKNRDPPLASKDA 240
OY 241 KROLGGAALGTHEDDKYRLDLAAGVGVVLDSSQGSIFQIMNKYIKDKYPMLOYI 300
    |||
DB 241 KROLGGAALGTHEDDKYRLDLAAGVGVVLDSSQGSIFQIMNKYIKDKYPMLOYI 300
OY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITQEVLAGCRPOATAVYKVEYARRFVGP 360
    |||
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITQEVLAGCRPOATAVYKVEYARRFVGP 360
OY 361 VIADGSIQVNGHIAKALALGASTVMMGSLAATTEARGEYFSDGIRLKKYRGMSLDM 420
    |||
DB 361 VIADGSIQVNGHIAKALALGASTVMMGSLAATTEARGEYFSDGIRLKKYRGMSLDM 420
OY 421 DKHLSSQNRYSFSEADKIKVAQGVSGAVDPDKGSIHKFVPLYLAGIQHSCODIGAKSLTOVR 480
    |||
DB 421 DKHLSSQNRYSFSEADKIKVAQGVSGAVDPDKGSIHKFVPLYLAGIQHSCODIGAKSLTOVR 480
OY 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEYKRLF 514
    |||
DB 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEYKRLF 514

RESULT 6
AAE18188
ID AAE18188 standard; protein; 514 AA.
XX
AC AAE18188;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human wild-type inosine 5'-monophosphate dehydrogenase (IMPDH).
XX
KW Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;
KW proliferative-type disease; cancer; imaging methodology; cytostatic;
KW enzyme; therapy.
XX
OS Homo sapiens.
XX

```





Best Local Similarity 83.3%; Pred. No. 2,2e-209;  
Matches 428; Conservative 39; Mismatches 47; Indels 0; Gaps 0;

```

OY 1 MADYLLISGTSVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFTADVDLTSALTKKIT 60
DB 1 MADYLLISGTSVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFTADVDLTSALTKKIT 60
OY 61 LKTPPLVSSPMDVTEADMAIAMAALMGIGFIHNCNCPREFQANVRKVKKFEQGITDPV 120
DB 61 LKTPPLVSSPMDVTEADMAIAMAALMGIGFIHNCNCPREFQANVRKVKKFEQGITDPV 120
OY 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVLCIISRDIIDFLKEEHDCELEIMT 180
DB 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVLCIISRDIIDFLKEEHDCELEIMT 180
OY 181 KREDIVVAPRSTTLKANEILORSKKGLPIVNEDELVAILARTDLKKNRDPPLASKDA 240
DB 181 KREDIVVAPRSTTLKANEILORSKKGLPIVNEDELVAILARTDLKKNRDPPLASKDA 240
OY 241 KKQLCGAAIGTHEDDKYRLDLAAGVVDVLDSSQGSNIFQINMKIKYIKDKYPMLOVY 300
DB 241 KKQLCGAAIGTHEDDKYRLDLAAGVVDVLDSSQGSNIFQINMKIKYIKDKYPMLOVY 300
OY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYVYEAARFVGP 360
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYVYEAARFVGP 360
OY 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEFFSDGRLKRYGMSLDM 420
DB 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEFFSDGRLKRYGMSLDM 420
OY 421 DKHLSSONRYFSEADKIKVAOGVSGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTOVR 480
DB 421 DKHLSSONRYFSEADKIKVAOGVSGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTOVR 480
OY 481 AMMYSGELKFEKRTSSAQVEGCVHLSHSEKRLF 514
DB 481 AMMYSGELKFEKRTSSAQVEGCVHLSHSEKRLF 514
OY 481 SMWYSGELKFEKRTSSAQVEGCVHLSHSEKRLY 514
DB 481 SMWYSGELKFEKRTSSAQVEGCVHLSHSEKRLY 514

RESULT 8
AAE18258
ID AAE18258 standard. Protein: 514 AA.
XX
AC AAE18258;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human type I IMPDH mutant, N109K.
XX
KW Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;
KW proliferative-type disease; cancer; imaging methodology; cytostatic;
KW enzyme; therapy; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 109 /note= "Wild-type Asn is substituted with Lys"
PI
XX WO200185952-A2.
XX
XX 15-NOV-2001.
XX
XX 10-MAY-2001: 2001WO-US15457.
XX
XX 10-MAY-2000: 2000US-203448P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;
XX Moutavieff JE, Einspahr HM, Kish K;
PI

```

XX WPI: 2002-164105/21.  
DR  
XX  
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
XX Disclosure: Page -: 161pp; English.

The invention relates to modified inosine 5'-monophosphate dehydrogenase (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain substituted for a subdomain of a wild-type IMPDH polypeptide. The modified IMPDH polypeptides are useful for drug discovery, for therapeutic, diagnostic and prognostic procedures for detecting or quantifying modified IMPDH polypeptides and their corresponding nucleic acids. IMPDH polypeptides are also useful for generating antibodies, as diagnostic and prognostic markers of diseases, as targets for various therapeutic modalities, and to identify and isolate ligands and other agents that bind to modified IMPDH. These antibodies may be used in diagnostic assays, imaging methodologies, therapeutic methods in the management of cancer or other proliferative-type diseases, and in purifying modified IMPDH polypeptides and for isolating related molecules such as wild type and mutant IMPDH polypeptides. The present sequence is human type I IMPDH mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on the human wild-type type I IMPDH  
CC SEQ.ID.NO.48 (AAE18185) shown in fig 1 of the specification.

XX Sequence 514 AA:

Query Match 85.0%; Score 2227; DB 23; Length 514;  
Best Local Similarity 83.3%; Pred. No. 3,4e-209;  
Matches 428; Conservative 39; Mismatches 47; Indels 0; Gaps 0;

```

OY 1 MADYLLISGTSVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFTADVDLTSALTKKIT 60
DB 1 MADYLLISGTSVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFTADVDLTSALTKKIT 60
OY 61 LKTPPLVSSPMDVTEADMAIAMAALMGIGFIHNCNCPREFQANVRKVKKFEQGITDPV 120
DB 61 LKTPPLVSSPMDVTEADMAIAMAALMGIGFIHNCNCPREFQANVRKVKKFEQGITDPV 120
OY 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVLCIISRDIIDFLKEEHDCELEIMT 180
DB 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVLCIISRDIIDFLKEEHDCELEIMT 180
OY 181 KREDIVVAPRSTTLKANEILORSKKGLPIVNEDELVAILARTDLKKNRDPPLASKDA 240
DB 181 KREDIVVAPRSTTLKANEILORSKKGLPIVNEDELVAILARTDLKKNRDPPLASKDA 240
OY 241 KKQLCGAAIGTHEDDKYRLDLAAGVVDVLDSSQGSNIFQINMKIKYIKDKYPMLOVY 300
DB 241 KKQLCGAAIGTHEDDKYRLDLAAGVVDVLDSSQGSNIFQINMKIKYIKDKYPMLOVY 300
OY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYVYEAARFVGP 360
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYVYEAARFVGP 360
OY 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEFFSDGRLKRYGMSLDM 420
DB 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEFFSDGRLKRYGMSLDM 420
OY 421 DKHLSSONRYFSEADKIKVAOGVSGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTOVR 480
DB 421 DKHLSSONRYFSEADKIKVAOGVSGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTOVR 480
OY 481 AMMYSGELKFEKRTSSAQVEGCVHLSHSEKRLF 514
DB 481 AMMYSGELKFEKRTSSAQVEGCVHLSHSEKRLF 514
OY 481 SMWYSGELKFEKRTSSAQVEGCVHLSHSEKRLY 514
DB 481 SMWYSGELKFEKRTSSAQVEGCVHLSHSEKRLY 514

RESULT 9

```

Thu Feb 13 13:18:45 2003

us-09-846-637c-4.rag

Page 7

```

AAE18185
ID AAE18185 standard: Protein: 514 AA.
XX
AC AAE18185:
XX
XX
DT 07-MAY-2002 (first entry)
XX
DE Human wild-type, type I IMPDH #1.
XX
DE Human: inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;
KM proliferative-type disease; cancer; imaging methodology; cytostatic;
KM enzyme; therapy.
XX
OS Homo sapiens.
XX
PN WO200185952-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001: 2001WO-US15457.
XX
PR 10-MAY-2000: 2000US-203448P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N:
PI Mouravieff JE, Einspahr HM, Kish K:
XX
DR WPI: 2002-164105/21.
XX
XX
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
PT having an oligo-peptide domain substituted for a subdomain of a
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating
XX antibodies.
XX
PS Claim 55: Fig 1: 161pp; English.
XX
XX
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The
CC modified IMPDH polypeptides are useful for drug discovery, for
CC therapeutic, diagnostic and prognostic procedures for detecting or
CC quantifying modified IMPDH polypeptides and their corresponding nucleic
CC acids. IMPDH polypeptides are also useful for generating antibodies, as
CC diagnostic and prognostic markers of diseases, as targets for various
CC therapeutic modalities, and to identify and isolate ligands and other
CC agents that bind to modified IMPDH. These antibodies may be used in
CC diagnostic assays, imaging methodologies, therapeutic methods in the
CC management of cancer or other proliferative-type diseases, and in
CC purifying modified IMPDH polypeptides and for isolating related
CC molecules such as wild type and mutant IMPDH polypeptides. The present
CC sequence is human wild-type type I IMPDH.
XX
XX
SQ Sequence 514 AA:

```

```

Query Match      84.8%; Score 2222; DB 23: Length 514;
Best Local Similarity 83.1%; Pred. No. 1,1e-208;
Matches 427; Conservative 39; Mismatches 48; Incls 0; Gaps 0;

```

```

DB 181 PRELVVAPAGVTLKEANEILLQSKKKGLPIVNDCELVAILARTDLKKNRDYPLASKDS 240
QY 241 KQOLLCGAALGTHEDDKRYRLDLAAGVUVVVDSSQGNISFIQIMKIKYIKRKYPMLOYI 300
DB 241 OKOLLGGAAGVTRREDDKRYRLDLTQAGVDVIVLDSQGNISVQIAMVHYIKRKYPHLOYI 300
QY 301 GGNVVTAAOAKNLDAGVVALRVGMGSGSICITIOEVLACGRPOATVAVYVEYARRFGVP 360
DB 301 GGNVVTAAOAKNLDAGVGLRVGMGSGSICITIOEYMGCRPOGTAIVYVAETARRFGVP 360
QY 361 VIADGGIQNVGHIAMALALGASTVMKGSLLAATTEAPCEYFFSDGIRLKYRGMSLDM 420
DB 361 IADGCIQTVGVHVALALGASTVMKGSLLAATTEAPCEYFFSDGIRLKYRGMSLDM 420
QY 421 DKHLSSQNRYSFADRIKVAQGVSGAVODKGSIHKFPYPLINGIOHSCODIGAKSLTOVR 480
DB 421 EKSSSSQKRYFESEGDVKVKAQGVSGSIQDKSITQKFPVPLIAGIOHGCODIGARSLTVLR 480
QY 481 AMYSGELKFEKRTSSAQVEGVSLSHYEKRLF 514
DB 481 SMYSGELKFEKRTMSAQIEGCVHGLSHYEKRLY 514

RESULT 10
AAE18190
ID AAE18190 standard: Protein: 514 AA.
XX
AC AAE18190:
XX
XX
DT 07-MAY-2002 (first entry)
XX
DE Human wild-type, type I IMPDH #2.
XX
XX
KM Human: inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;
KM proliferative-type disease; cancer; imaging methodology; cytostatic;
KM enzyme; therapy.
XX
XX
OS Homo sapiens.
XX
PN WO200185952-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001: 2001WO-US15457.
XX
PR 10-MAY-2000: 2000US-203448P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N:
PI Mouravieff JE, Einspahr HM, Kish K:
XX
DR WPI: 2002-164105/21.
XX
XX
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
PT having an oligo-peptide domain substituted for a subdomain of a
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating
XX antibodies.
XX
PS Claim 55: Page 157-159; 161pp; English.
XX
XX
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The
CC modified IMPDH polypeptides are useful for drug discovery, for
CC therapeutic, diagnostic and prognostic procedures for detecting or
CC quantifying modified IMPDH polypeptides and their corresponding nucleic
CC acids. IMPDH polypeptides are also useful for generating antibodies, as
CC diagnostic and prognostic markers of diseases, as targets for various
CC therapeutic modalities, and to identify and isolate ligands and other
CC agents that bind to modified IMPDH. These antibodies may be used in
CC diagnostic assays, imaging methodologies, therapeutic methods in the
CC management of cancer or other proliferative-type diseases, and in

```

CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human wild-type type I IMPDH.

XX Sequence 514 AA:

Query Match 84.1%; Score 2202; DB 23; Length 514;  
Best Local Similarity 82.1%; Pred. No. 9,8e-207;  
Matches 422; Conservative 41; Mismatches 51; Indels 0; Gaps 0;

```
OY 1 MADYLIISGTSVPPDGLTAQDLFNCGDLTYNDLFLPGYIDFADQVDTLSATKKT 60
DB 1 MADYLIISGTSVPPDGLTAQDLFNCGDLTYNDLFLPGYIDFADQVDTLSATKKT 60
OY 61 LKTPLVSSMDVTVTAGMAIMALMGIGIFHHNCTPEFOANEVKKVKYFEGGFTTDPV 120
DB 61 LKTPLVSSMDVTVTAGMAIMALMGIGIFHHNCTPEFOANEVKKVKYFEGGFTTDPV 120
OY 121 LSPKDRVDVFEAKARHGFQGIPTDTGRMSRLVGISSRDIDFLKEEHDCFLLEIMT 180
DB 121 LSPKDRVDVFEAKARHGFQGIPTDTGRMSRLVGISSRDIDFLKEEHDCFLLEIMT 180
OY 181 KREDLVNAPRSTTLKEANIIQRSKKKRLPIYNEDDELVAIIARTDKKNRPYPLASKA 240
DB 181 PRIELVAPAGVTLKEANEIIQRTKKGKLPYVNDDELVAIIARTDKKNRPYPLASKS 240
OY 241 KCOLLGAAGTGHEDDKYRLDLAAGVVDVVLDSOGNSIFOIMNKYIKRKYFNLQVY 300
DB 241 OKOLLCGAAGVTRDDKYRLDLTQAGVVDVYFHSOGNSVQIAMVHIKRYHQLQVY 300
OY 301 GGNVYTAQAQAKMLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYKVEYARRFGVP 360
DB 301 GGNVYTAQAQAKMLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYKVEYARRFGVP 360
OY 361 VIADGCIQVGHVIAKALALGASTVMGSLAATTAPEGFYFSDGIRLKKYGMGSLDM 420
DB 361 VIADGCIQVGHVIAKALALGASTVMGSLAATTAPEGFYFSDGIRLKKYGMGSLDM 420
OY 421 DKHLSSONRYFESEADKIKYAGCVSAGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTQVR 480
DB 421 EKSSSSQKRYFESEADKIKYAGCVSAGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTQVR 480
OY 481 AMMYGELKFEKRTSSADEVGVSHLSHYEKRLE 514
DB 481 SMYSGELKFEKRTSSADEVGVSHLSHYEKRLE 514
```

RESULT 11  
AAE18191  
ID AAE18191 standard; Protein: 514 AA.  
XX AAE18191;  
AC AAE18191;  
XX 07-MAY-2002 (first entry)  
DT 07-MAY-2002 (first entry)  
XX Human wild-type, type I IMPDH #3.  
DE Human wild-type, type I IMPDH #3.  
XX Human: inosine 5'-monophosphate dehydrogenase; IMPDH: drug discovery;  
KW proliferative-type disease; cancer; imaging methodology; cytostatic;  
XX enzyme; therapy.  
OS Homo sapiens.  
XX WO200185952-A2.  
PN 15-NOV-2001.  
XX 10-MAY-2001; 2001WO-US15457.  
XX 10-MAY-2000; 2000US-203448P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX

PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff JE, Einspahr HM, Kish K;  
XX WPI: 2002-164105/21.  
DR

XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX

PS Claim 55; Page 159-161; 161pp; English.

XX The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human wild-type type I IMPDH.

XX Sequence 514 AA:

Query Match 84.0%; Score 2201; DB 23; Length 514;  
Best Local Similarity 82.3%; Pred. No. 1.2e-206;  
Matches 423; Conservative 39; Mismatches 52; Indels 0; Gaps 0;

```
OY 1 MADYLIISGTSVPPDGLTAQDLFNCGDLTYNDLFLPGYIDFADQVDTLSATKKT 60
DB 1 MADYLIISGTSVPPDGLTAQDLFNCGDLTYNDLFLPGYIDFADQVDTLSATKKT 60
OY 61 LKTPLVSSMDVTVTAGMAIMALMGIGIFHHNCTPEFOANEVKKVKYFEGGFTTDPV 120
DB 61 LKTPLVSSMDVTVTAGMAIMALMGIGIFHHNCTPEFOANEVKKVKYFEGGFTTDPV 120
OY 121 LSPKDRVDVFEAKARHGFQGIPTDTGRMSRLVGISSRDIDFLKEEHDCFLLEIMT 180
DB 121 LSPKDRVDVFEAKARHGFQGIPTDTGRMSRLVGISSRDIDFLKEEHDCFLLEIMT 180
OY 181 KREDLVNAPRSTTLKEANIIQRSKKKRLPIYNEDDELVAIIARTDKKNRPYPLASKA 240
DB 181 PRIELVAPAGVTLKEANEIIQRTKKGKLPYVNDDELVAIIARTDKKNRPYPLASKS 240
OY 241 KCOLLGAAGTGHEDDKYRLDLAAGVVDVVLDSOGNSIFOIMNKYIKRKYFNLQVY 300
DB 241 OKOLLCGAAGVTRDDKYRLDLTQAGVVDVYFHSOGNSVQIAMVHIKRYHQLQVY 300
OY 301 GGNVYTAQAQAKMLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYKVEYARRFGVP 360
DB 301 GGNVYTAQAQAKMLIDAGVDALRVGMGSGICITIOEVLACGRPOATAVYKVEYARRFGVP 360
OY 361 VIADGCIQVGHVIAKALALGASTVMGSLAATTAPEGFYFSDGIRLKKYGMGSLDM 420
DB 361 VIADGCIQVGHVIAKALALGASTVMGSLAATTAPEGFYFSDGIRLKKYGMGSLDM 420
OY 421 DKHLSSONRYFESEADKIKYAGCVSAGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTQVR 480
DB 421 EKSSSSQKRYFESEADKIKYAGCVSAGAVODKGSIHKEFVPLIAGIOHSCODIGAKSLTQVR 480
OY 481 AMMYGELKFEKRTSSADEVGVSHLSHYEKRLE 514
DB 481 SMYSGELKFEKRTSSADEVGVSHLSHYEKRLE 514
```

RESULT 12  
ABG22269

ID ABG22269 standard: Protein: 604 AA.  
XX  
AC ABG22269;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #22260.  
XX  
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001: 2001WO-US08631.  
XX  
PR 31-MAR-2000: 2000US-0540217.  
PR 23-AUG-2000: 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
DR N-PSDB: AAS86456.  
XX  
XX New isolated polynucleotide and encoded polypeptides useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20: SEQ ID NO 52628; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridization probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations in  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 604 AA:

Query Match 81.8%; Score 2141.5; DB 22: length 604;  
Best Local Similarity 80.2%; Pred. No. 11e-200;  
Matches 413; Conservative 43; Mismatches 58; Gaps 1; Gaps 1;

QY 1 MADYLLISGCTSYVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFADYVDLTSALTAKKIT 60  
DB 90 MADYLLISGCTSYVPPDGLTAQOLFNCGDGLTYNDLFLPGYIDFADYVDLTSALTAKKIT 149  
QY 61 LKTPPLVSSMDVVTFAFGMAIAMLTCGTFIHNCNPFPEOANVVRKVKYEGCFITDPVY 120  
DB 150 LKTPPLVSSMDVVTFAFGMAIAMLTCGTFIHNCNPFPEOANVVRKVKYEGCFITDPVY 209  
QY 121 LSPKRDVRDVFCAKARHGFCGIPITDTGRMGSLVGISSRDIDFLKFEHDCFLIEIMT 180

DB 210 LSPSHTVGVLDLECKMRHSGCIPITETGTMGSKLVGIVTSRDIDFLAEKDHTTLSEVMT 269  
QY 181 KREDLVVAPRSITLKEANETLQSKKGLPIYNEDELVAIARTDLKKNRDYPLASKDA 240  
DB 270 PRIELVYAPAGVTLKEBANETLQSKKGLPIYNDCEDEVAIARTDLKKS RDPYPLASKDS 329  
QY 241 KROLGCAALGTHEDDKRYRLDILAQAGVYVYLDSSQGSIPQINMKIKYIKQKYPPLQYI 300  
DB 330 QPOLLCGAALVREEDDKRYRLDILTQAGVYVYLDSSQGSVYQYARVLPKQYPPLOLYL 389  
QY 301 GGNVVTAAQAKNLDIAGVDALRVGMGSGSICIIOEYLAACRPQAT-AVYKYVEYARRFCV 359  
DB 390 GGNVVTAAQAKNLDIAGVDGLRVGMCGSICITQEVYMACGRPGQCCVTRXAEYARRFCV 449  
QY 360 PYIADGCIQNGHIAKALALGASTVVMGSLAATTAPPEYFFSDSIRLKKTRGMSLDA 419  
DB 450 PITADGCIQNGHIAKALALGASTVVMGSLAATTAPPEYFFSDSIRLKKTRGMSLDA 509  
QY 420 MDKHLSSQNRPFSEADKIRVAVQSGAVODKGSIHKFPYLLAGI0HSCODIGAKSLTQV 479  
DB 510 MEKSSSSQKRYFSEGDQKVAIAQCVSSGSIQDKGSIQKFVYLLAGI0HSCODIGAKSLTQV 569  
QY 480 RAMMYSGLKFEKRRTSSAQVEGCVHSLHSEYKRLF 514  
DB 570 RSMYSGELKFEKRRTMSPQIEGVHGLHSEYKRLY 604

RESULT 13  
AAU99361  
ID AAU99361 standard: Protein: 509 AA.  
XX  
XX AAU99361;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
XX Human oxidoreductase 1 (OXR1) protein.  
DE  
XX  
XX Human: enzyme: oxidoreductase 1; OXR1; oxidation-reduction;  
KW aerobic metabolism; free radical; superoxide; hydroxyl ion; oxidation;  
KW induction; oxidation potential; gene therapy; rheumatoid arthritis;  
KW inflammatory bowel disease; cell proliferation; psoriasis; lymphoma;  
KW cancer; Parkinson's disease; atherosclerosis; neurological disorder;  
KW epilepsy; Alzheimer's disease; myopathy; masthenia gravis;  
KW periodic paralysis; mental disorder; SAD; seasonal affective disorder;  
KW amnesia; metabolic disorder; angina; Down's syndrome; muscular dystrophy;  
KW smooth muscle disorder; asthma; hypertension; incyte 72101265CD1.  
XX  
XX Homo sapiens.  
OS  
PN WO200250284-A2.  
XX  
PD 27-JUN-2002.  
XX  
XX 18-DEC-2001: 2001WO-US49131.  
PF 21-DEC-2000: 2000US-257802P.  
PR 18-JAN-2001: 2001US-262901P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Tribouley CM, Lee EA, Yao MG, Elliott VS, Yue H;  
PI WPI: 2002-528450/56.  
DR N-PSDB: ABR89032.  
XX  
XX New human oxidoreductase polypeptides and polynucleotides useful for  
PT diagnosing, treating and preventing e.g. cell proliferative (e.g.  
PT arteriosclerosis), neurological (e.g. epilepsy), metabolic (e.g.  
PT hyperproliferemia) -  
XX  
XX Claim 1: Page 113-114; 119pp; English.

The invention discloses isolated human polypeptides, and encoding polypeptides, comprising the oxidoreductases, OXR-1, OXR-2 and OXR-3. Eukaryotic cells extract energy and synthesise macromolecules by a complex series of oxidation-reduction reactions, collectively called aerobic metabolism. A consequence of aerobic metabolism is the production of free radicals in the form of superoxides and hydroxyl ions. Defects in enzymes involved in oxidation and reduction reactions in cells (oxidoreductases) lead to imbalances in the oxidation potential within cells and frequently with clinical manifestations. The polypeptides and polynucleotides can be used for screening compounds for agonists or antagonists, for screening for compounds that specifically bind to them, for screening for compounds that modulates their activity, for screening compounds which alter expression of a target polynucleotide, for diagnosing conditions or diseases associated with the expression of OXR in a subject, for raising antibodies and in gene therapy. The nucleic acid and amino acid sequences are useful in the diagnosis, treatment and prevention of rheumatoid arthritis, inflammatory bowel disease, cell proliferative disorders (e.g. atherosclerosis, psoriasis, lymphoma and cancers), neurological disorders (e.g. epilepsy, Alzheimer's disease and Parkinson's disease), myopathies, myasthenia gravis, periodic paralysis, mental disorders (e.g. seasonal affective disorder (SAD) and anorexia), metabolic disorders (e.g. Down's syndrome and muscular dystrophy) and smooth muscle disorders (e.g. angina, asthma and hypertension). The sequence presented is the human oxidoreductase 1 (OXR-1) protein (Incyte ID No: 72101285CD1).

SQ Sequence 509 AA;

Query Match	80.88; Score 2115.5; DB 23; Length 509;
-------------	---

Best Local Similarity 79.8%; Pred. No. 2.9e-198;

Matches 410; Conservative 44; Mismatches 55; Indels 5; Gaps 1;

QY	1	MADYIISGTSYVDDGJTAQOLPNCBGDLTYNDFLLPGYIDFTADQVUDJTSALTKKIT	60
Db	1	MADYIISGTSYVDDGJTAQOLPNCBGDLTYNDFLLPGYIDFTADQVUDJTSALTKKIT	60
QY	61	LKTPVSSPMQTVJEAGAIMALTGGIGFIHNHCPEFQANVEKRYKXKEGSEITDPVY	120
Db	61	LKTPVSSPMQTVJEAGAIMALTGGIGFIHNHCPEFQANVEKRYKXKEGSEITDPVY	120
QY	121	LSPKRDVDFEAKARHFGCIPITPDGRMGSRVJGISSRDIIDFLKEEHEDFLEIIMT	180
Db	121	LSPKRDVDFEAKARHFGCIPITPDGRMGSRVJGISSRDIIDFLKEEHEDFLEIIMT	180
QY	181	KREDLVAPRSITLKEANEILQORSKCKPLPVNDEDELVALIARTDLKKNRDYPLASKDA	240
Db	181	KREDLVAPRSITLKEANEILQORSKCKPLPVNDEDELVALIARTDLKKNRDYPLASKDA	240
QY	241	KKOLLGCAITGTHEDDKTRDLLAQAGVDVYVLDSOGNSLFOJNMKITKDKYPNLQYI	300
Db	241	KKOLLGCAITGTHEDDKTRDLLAQAGVDVYVLDSOGNSLFOJNMKITKDKYPNLQYI	300
QY	301	GCNVYTTAAQANLLIDAGVDALRVBMGSGSICIIOEVLACRPOATAYKVEKXARRPGVP	360
Db	301	GCNVYTTAAQANLLIDAGVDALRVBMGSGSICIIOEVLACRPOATAYKVEKXARRPGVP	360
QY	361	VJADGCIQNVGHIIKALALGASTYMMGSSLAAATTEADGEFFSDGLKLYKRYRMGSLDM	420
Db	361	VJADGCIQNVGHIIKALALGASTYMMGSSLAAATTEADGEFFSDGLKLYKRYRMGSLDM	420
QY	421	DKHLSSQNRFRSEADKIKVAGSGAVQDKGSIHKKFVYPIIAGIQHSCODIGAKSLTOYA	480
Db	421	DKHLSSQNRFRSEADKIKVAGSGAVQDKGSIHKKFVYPIIAGIQHSCODIGAKSLTOYA	480
QY	481	AMNYSGELKFEKRTISSADEVGCVSLHSYERLRF	514
Db	481	AMNYSGELKFEKRTISSADEVGCVSLHSYERLRF	514
QY	514	SMYSGELKFEKRTISSADEVGCVSLHSYERLRF	548
Db	514	SMYSGELKFEKRTISSADEVGCVSLHSYERLRF	548

AC	AAEL8165,	
XX		
DT	07-MAY-2002	(first entry)
XX		
DE	Human type II inosine 5'-monophosphate dehydrogenase (IMPDH)-DKT variant.	
XX		
KM	Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;	
KW	proliferative-type disease; cancer; imaging methodology; cytostatic;	
XX	therapy; enzyme; variant.	
OS	Homo sapiens.	
XX	Synthetic.	
EH		
FT	Key	Location/Qualifiers
FT	Domain	1..110
FT		/note= "N-terminal catalytic core domain"
FT	Domain	114..384
FT		/note= "C-terminal catalytic core domain"
XX		
PN	WO200185952-A2.	
PD		
PD	15-NOV-2001.	
XX		
PF	10-MAY-2001; 2001WO-US15457.	
XX		
PR	10-MAY-2000; 2000US-203448P.	
XX		
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.	
PI	Krystek SR, Sheriff S, Witmer MR, Hollenbaugh DL, Yan N;	
PI	Mouravieff JE, Einspahr HM, Kish K;	
XX		
DR	WPI: 2002-164105/21.	
XX	N-PSDB: AAD28922.	
PT	New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide	
PT	having an oligo-peptide domain substituted for a subdomain of a	
PT	wild-type IMPDH polypeptide, useful in drug discovery or for generating	
PT	antibodies -	
XX		
PS	Claim 8; Fig 4; 161pp; English.	
XX		
CC	The invention relates to modified inosine 5'-monophosphate dehydrogenase	
CC	(IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain	
CC	substituted for a subdomain of a wild-type IMPDH polypeptide. The	
CC	modified IMPDH polypeptides are useful for drug discovery, for	
CC	therapeutic, diagnostic and prognostic procedures for detecting or	
CC	quantifying modified IMPDH polypeptides and their corresponding nucleic	
CC	acids. IMPDH polypeptides are also useful for generating antibodies, as	
CC	diagnostic and prognostic markers of diseases, as targets for various	
CC	therapeutic modalities, and to identify and isolate ligands and other	
CC	agents that bind to modified IMPDH. These antipodes may be used in	
CC	management of cancer or other proliferative-type diseases, and in	
CC	purifying modified IMPDH polypeptides and for isolating related	
CC	molecules such as wild type and mutant IMPDH polypeptides. The present	
CC	sequence is human type II IMPDH-DKT variant. This sequence comprises	
XX	substitute tri-peptide DKT for the domain region of wild-type IMPDH.	
XX	Sequence	384 AA:

SQ Sequence 384 AA;

Query Match	71.18;	Score 1862;	DB 23;	Length 384;
-------------	--------	-------------	--------	-------------

Best Local Similarity 73.98; Pred. No. 1.4e-173;

Matches 380; Conservative 1; Mismatches 3; Indels 130; Gaps 1;

```

0y 1 MADYLIISGTSYVDPDGLTAQQLFENGCGGLTYNDPFLPGIIPDTAAQVLDLSALTKRIT 60
      |||||
Db 1 MADYLIISGTSYVDPDGLTAQQLFENGCGGLTYNDPFLPGIIPDTAAQVLDLSALTKRIT 60
      |||||
0y 61 LKTPLVSSPMDTVTEAGMAIMALTGSGIGFIHNHCTEPEQANERKVKRYKEQGITDPVY 120
      |||||
Db 61 LKTPLVSSPMDTVTEAGMAIMALTGSGIGFIHNHCTEPEQANERKVKRYKEQGITDPVY 120
      |||||

```

QY 121 LSPKDRVDVFEAKARRHFGCIPITDTGRMGSRVGLISSRDIDFLKEEHDCFLEEIMT 180  
Db 112 ----- 111  
QY 181 KREDLVVAPRSTILKEANELIORSKKKGLPIVNEDELVAIARTDLKKNRDYPLASKDA 240  
Db 112 ----- 111  
QY 241 KQOLCGAATGTHEDDKYRLDLLAAGVADVVDLSSOGNSIFOINMKTIKDKYPLQVY 300  
Db 112 KTLGGAATGTHEDDKYRLDLLAAGVADVVDLSSOGNSIFOINMKTIKDKYPLQVY 170  
QY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITIOEVLACGRPOATAVYKVEYARRFGVP 360  
Db 171 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITIOEVLACGRPOATAVYKVEYARRFGVP 230  
QY 361 VIADGGTQNVGHTAKALALAGASTVMMGSLAATTEAPGEYFFSDGIRLKYRGMGSLDAM 420  
Db 231 VIADGGTQNVGHTAKALALAGASTVMMGSLAATTEAPGEYFFSDGIRLKYRGMGSLDAM 290  
QY 421 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYIAGIQHSCDDIGAKSLTOVR 480  
Db 291 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYIAGIQHSCDDIGAKSLTOVR 350  
QY 481 AMMYSGELKFEKRTSSAOVEGVHSLHSYEKRLF 514  
Db 351 AMMYSGELKFEKRTSSAOVEGVHSLHSYEKRLF 384  
RESULT 15  
AAE18179  
ID AAE18179 standard: Protein; 385 AA.  
XX  
AC AAE18179;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Human modified inosine 5'-monophosphate dehydrogenase (IMPDH) #10.  
XX  
KM Human: inosine 5'-monophosphate dehydrogenase: IMPDH; drug discovery;  
KM proliferative-type disease; cancer; imaging methodology; cytostatic;  
KM therapy; enzyme.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN MO200185952-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 10-MAY-2001: 2001WO-US15457.  
XX  
PR 10-MAY-2000: 2000US-203448P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Krystek SR, Sherliff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff JE, Einspahr HM, Kish K;  
XX  
DR WPI: 2002-164105/21.  
XX  
PS Claim 8: Page 130-131: 161pp: English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.203) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or

CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human modified IMPDH. This sequence comprises a substitute  
CC tetra-peptide for the subdomain region of wild-type IMPDH.  
XX  
SQ Sequence 385 AA:  
Query Match 71.1%; Score 1861.5; DB 23; Length 385;  
Best Local Similarity 74.3%; Pred. No. 1.5e-173;  
Matches 382; Conservative 0; Mismatches 3; Indels 129; Gaps 2;  
QY 1 MADYLISSGTSYPPDGLTAQQLFNCGGDGLTYNDFLLPEYIDFTADQVDLTALTKKIT 60  
Db 1 MADYLISSGTSYPPDGLTAQQLFNCGGDGLTYNDFLLPEYIDFTADQVDLTALTKKIT 60  
QY 61 LKTPVSSPMDVTYTAGMATAMALTGIGIFIHNCCTPEFOANPVRKYKKEGFTIDPVV 120  
Db 61 LKTPVSSPMDVTYTAGMATAMALTGIGIFIHNCCTPEFOANPVRKYKKEGFTIDPVV 110  
QY 121 LSPKDRVDVFEAKARRHFGCIPITDTGRMGSRVGLISSRDIDFLKEEHDCFLEEIMT 180  
Db 111 -SP----- 112  
QY 181 KREDLVVAPRSTILKEANELIORSKKKGLPIVNEDELVAIARTDLKKNRDYPLASKDA 240  
Db 113 ----- 112  
QY 241 KQOLCGAATGTHEDDKYRLDLLAAGVADVVDLSSOGNSIFOINMKTIKDKYPLQVY 300  
Db 113 -TOLLGGAATGTHEDDKYRLDLLAAGVADVVDLSSOGNSIFOINMKTIKDKYPLQVY 171  
QY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITIOEVLACGRPOATAVYKVEYARRFGVP 360  
Db 172 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICITIOEVLACGRPOATAVYKVEYARRFGVP 231  
QY 361 VIADGGTQNVGHTAKALALAGASTVMMGSLAATTEAPGEYFFSDGIRLKYRGMGSLDAM 420  
Db 232 VIADGGTQNVGHTAKALALAGASTVMMGSLAATTEAPGEYFFSDGIRLKYRGMGSLDAM 291  
QY 421 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYIAGIQHSCDDIGAKSLTOVR 480  
Db 292 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYIAGIQHSCDDIGAKSLTOVR 351  
QY 481 AMMYSGELKFEKRTSSAOVEGVHSLHSYEKRLF 514  
Db 352 AMMYSGELKFEKRTSSAOVEGVHSLHSYEKRLF 385

Search completed: February 13, 2003, 04:50:19  
Job time : 89 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 15:12:14 : Search time: 203 Seconds  
(without alignments)  
11452.782 Million cell updates/sec

Title: US-09-846-637C-3

Perfect score: 1654

Sequence: 1 gaatcggcggtcctcgga.....gttagaagaccgcgaattc 1654

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vl:\*  
30: em.htg.hum:\*  
31: em.htg.inv:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.pod:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1647.6	99.6	1654	9 HUMIMP	J04208 Human inosi
2	1635.6	98.9	1642	6 AR141869	AR141869 Sequence
3	1635.6	98.9	1642	6 I06570	I06570 Sequence 1
4	1622.6	98.1	1663	9 BC006124	BC006124 Homo sapi
5	1616.6	97.7	1648	9 BC012840	BC012840 Homo sapi
6	1614.6	97.6	1655	9 BC015567	BC015567 Homo sapi
7	1337	80.2	225727	2 AC122514	AC122514 Mus muscu
8	1326.6	79.3	1707	10 MUSIMPDA	M39334 Mouse IMP d
9	1322.2	79.9	1680	10 BC010314	BC010314 Mus muscu
10	1321.8	79.9	1620	6 AR141890	AR141890 Sequence
11	1321.8	79.9	1620	6 I06572	I06572 Sequence 3
12	1320.6	79.8	1614	10 CRUIMP	J04209 Chinese ham
13	1298.6	78.5	1545	10 MUSIMPDA	M98333 Mus musculu
14	963.6	58.3	2479	9 AK054667	AK054667 Homo sapi
15	963.6	58.3	2520	9 BC033622	BC033622 Homo sapi
16	963.6	58.3	2547	9 AK092452	AK092452 Homo sapi
17	960.8	58.1	2858	6 AX409472	AX409472 Sequence
18	960.8	58.1	2858	6 HUMIMP	J05272 Human IMP d
19	937	56.7	1545	10 NMU00978	U00978 Mus musculu
20	906.4	54.8	172152	9 AC091807	AC091807 Homo sapi
21	906.4	54.8	175312	2 AC027551	AC027551 Homo sapi
22	906.4	54.8	176295	9 AC069362	AC069362 Homo sapi
23	843.2	51.0	48537	9 AL450306	AL450306 Human DNA
24	841.6	50.5	170537	9 AC008265	AC008265 Homo sapi
25	841.6	50.5	184425	9 AC025593	AC025593 Homo sapi
26	839	50.7	151990	9 AL354989	AL354989 Human DNA
27	829.2	50.1	2277	9 AK054640	AK054640 Homo sapi
28	820	49.6	164916	9 AC021305	AC021305 Homo sapi
29	820	49.6	166164	2 AC023188	AC023188 Homo sapi
30	820	49.6	178771	2 AF267168	AF267168 Homo sapi
31	820	49.6	197310	2 AF254981	AF254981 Homo sapi
32	816.4	49.4	1158	6 AX418035	AX418035 Sequence
33	813.2	49.2	1158	6 AX418036	AX418036 Sequence
34	810	49.0	1158	6 AX418037	AX418037 Sequence
35	809.6	48.9	1155	6 AX418030	AX418030 Sequence
36	809.6	48.9	1155	6 AX418031	AX418031 Sequence
37	809.6	48.9	1155	6 AX418032	AX418032 Sequence
38	809.6	48.9	1155	6 AX418033	AX418033 Sequence
39	806.8	48.8	168608	2 AC103220	AC103220 Rattus no
40	799.6	48.3	160285	9 AC090945	AC090945 Homo sapi
41	793.2	48.0	144848	2 AL592165	AL592165 Homo sapi
C 42	771.6	46.7	208924	9 AC018735	AC018735 Homo sapi
C 43	769.2	46.5	8801	9 AB038965S4	AB038968 Homo sapi
44	769.2	46.5	127361	9 AC007283	AC007283 Homo sapi
C 45	761	46.0	172237	2 AC131389	AC131389 Homo sapi

#### ALIGNMENTS

RESULT 1  
HUMIMP  
LOCUS HUMIMP 1654 bp mRNA linear PRI 11-JUN-1993  
DEFINITION Human inosine-5'-monophosphate dehydrogenase (IMP) mRNA, complete cds.  
ACCESSION J04208  
VERSION J04208.1 GI:186391  
KEYWORDS inosine-5'-monophosphate dehydrogenase.  
SOURCE Human peripheral blood leukocyte, CDNA to mRNA, clone HUMIMP.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1654)  
Collart,F.R. and Huberman,E.  
Cloning and sequence analysis of the human and Chinese hamster

Pred. No. is the number of results predicted by chance to have a

JOURNAL J. Biol. Chem. 263 (30), 15769-15772 (1988)  
MEDLINE 89008491  
PUBMED 2902093

COMMENT Draft entry and computer-readable sequence for [1] kindly provided  
by F.R.Collart, 22-AUG-1988.

FEATURES  
Location/Qualifiers

CDS

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
48..1592  
/note="Inosine-5'-monophosphate dehydrogenase (EC  
1.1.1.205)"  
/codon\_start=1  
/protein\_id="AAA36112.1"  
/db\_xref="GI:307066"

BASE COUNT 415 a 418 c 451 g 370 t  
ORIGIN 45 bp upstream of Balli site.

Query Match 99.6% Score 1647.6; DB 9; Length 1654;  
Best Local Similarity 99.8% Pred. No. 0;  
Matches 1650; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GAATTCGGCGGCGCTCGAGACACCGCGGTCCTGTTGGCCATGGCCGACTACC 60  
DB 1 GAATTCGGCGGCTCTCGAGACACCGCGGCTCTCTGTTGGCCATGGCCGACTACC 60  
OY 61 TGATTAGTGGGGGACGTCCTACGTGCTCCAGACGACGACTCACAGCAGCAGCAAGCTTTGA 120  
DB 61 TGATTAGTGGGGGACGTCCTACGTGCTCCAGACGACGACTCACAGCAGCAGCAAGCTTTGA 120  
OY 121 ACTCGGAGACGCGCTCACCTACATGACTTTCTCTCCCTGGGTACATGACTTCA 180  
DB 121 ACTCGGAGACGCGCTCACCTACATGACTTTCTCTCCCTGGGTACATGACTTCA 180  
OY 121 ACTCGGAGACGCGCTCACCTACATGACTTTCTCTCCCTGGGTACATGACTTCA 180  
DB 121 ACTCGGAGACGCGCTCACCTACATGACTTTCTCTCCCTGGGTACATGACTTCA 180  
OY 181 CTGCACACGAGGTGACCTGACTTGTGCTGACCAAGAAATCACTTTAAGACCCGAC 240  
DB 181 CTGCACACGAGGTGACCTGACTTGTGCTGACCAAGAAATCACTTTAAGACCCGAC 240  
OY 241 TGGTTCTCTCCATGACACAGTACACAGGCTGGATGGCCATAGCAATGGCGCTTA 300  
DB 241 TGGTTCTCTCCATGACACAGTACACAGGCTGGATGGCCATAGCAATGGCGCTTA 300  
OY 301 CAGGCGGTATTGGCTTCATCCACCACTGTAACCTGTAATTCAGGCCAATGAAGTTT 360  
DB 301 CAGGCGGTATTGGCTTCATCCACCACTGTAACCTGTAATTCAGGCCAATGAAGTTT 360  
OY 361 GGAAGTGAAGAAATATGAACAGGATTCATCACAGACCTGTGGTCCAGCCCAAG 420  
DB 361 GGAAGTGAAGAAATATGAACAGGATTCATCACAGACCTGTGGTCCAGCCCAAG 420  
OY 421 ATCGCTCGCGGATGTTTGTGAGGCCAAGCCCGGATGGTTCTCGCGATGCCATCA 480  
DB 421 ATCGCTCGCGGATGTTTGTGAGGCCAAGCCCGGATGGTTCTCGCGATGCCATCA 480  
OY 481 CAGACACAGCGCGGATGGGAGCGCTTGGTGGCATCATCTCTCCAGGAGCATGGATT 540  
DB 481 CAGACACAGCGCGGATGGGAGCGCTTGGTGGCATCATCTCTCCAGGAGCATGGATT 540  
OY 541 TTCTCAAGAGGGAACATGACTTTCTTGGAAGAGATTAATGACAAAGGAGGAGACT 600  
DB 541 TTCTCAAGAGGGAACATGACTTTCTTGGAAGAGATTAATGACAAAGGAGGAGACT 600

OY 601 TGGTGTACCCCCCGACGATCACACTGAGAGGCAATGAATTTCTGACGCCACGA 660  
DB 601 TGGTGTAGCCCCCGACGATCACACTGAGAGGCAATGAATTTCTGACGCCACGA 660  
OY 661 AAGAGGAAAGTTGCCATTTGAATGAAGATGATGAGCTTGGCCATTCATGGCCGA 720  
DB 661 AAGAGGAAAGTTGCCATTTGAATGAAGATGATGAGCTTGGCCATTCATGGCCGA 720  
OY 721 CAGACCTGAAGAAATCGGAGTACCCATAGCTTCGCAAGATGCCAAGAAAGCTGTC 780  
DB 721 CAGACCTGAAGAAATCGGAGTACCCATAGCTTCGCAAGATGCCAAGAAAGCTGTC 780  
OY 781 TGTGTGGGGCAGCCATTGGGCACTATGAGATGACAAATATGAGCTGCTGCTGCC 840  
DB 781 TGTGTGGGGCAGCCATTGGGCACTATGAGATGACAAATATGAGCTGCTGCTGCC 840  
OY 841 AGGCTGTGTGATGATGAGTGTGGTGGACTTCTCCAGGAAATTCATCTTCAGATCA 900  
DB 841 AGGCTGTGTGATGATGAGTGTGGTGGACTTCTCCAGGAAATTCATCTTCAGATCA 900  
OY 901 ATATGATCAAGTACATCAAGACAAATACCTTAATCTCCAGTCAATTTGAGGCAATGTGG 960  
DB 901 ATATGATCAAGTACATCAAGACAAATACCTTAATCTCCAGTCAATTTGAGGCAATGTGG 960  
OY 961 TCACTGCTCCCAAGCCAAAGACCTCATTGATGACAGGTGTGATGCCCTGGCGGTGGCA 1020  
DB 961 TCACTGCTCCCAAGCCAAAGACCTCATTGATGACAGGTGTGATGCCCTGGCGGTGGCA 1020  
OY 1021 TGGGAAGTGGCTCCATCTGATTCATTCACAGAGTGTGGCTGTGGGGCGGCCCAAGCA 1080  
DB 1021 TGGGAAGTGGCTCCATCTGATTCATTCACAGAGTGTGGCTGTGGGGCGGCCCAAGCA 1080  
OY 1081 CAGCAGTGTACAGGTGTATGATGATGACAGGCGCTTTGGTTCGGATTTGCTGATG 1140  
DB 1081 CAGCAGTGTACAGGTGTATGATGATGACAGGCGCTTTGGTTCGGATTTGCTGATG 1140  
OY 1141 GAGGAATCCAAATGTGGTCAATATTGCGAAAGCTTGGCCCTTGAGGCTCCACAGTCA 1200  
DB 1141 GAGGAATCCAAATGTGGTCAATATTGCGAAAGCTTGGCCCTTGAGGCTCCACAGTCA 1200  
OY 1201 TGAATGGGCTCTCCCTGCGTGCACACCTGAGGCGCTGGTGAATCTTTTCCGATG 1260  
DB 1201 TGAATGGGCTCTCCCTGCGTGCACACCTGAGGCGCTGGTGAATCTTTTCCGATG 1260  
OY 1261 GGATCCGGCTAAAGAAATATCGCGGTATGGTTCCTGATGCCATGACAGACCTCA 1320  
DB 1261 GGATCCGGCTAAAGAAATATCGCGGTATGGTTCCTGATGCCATGACAGACCTCA 1320  
OY 1321 GCAGCCAGAACAGATATTTCAAGTGAAGCTGACAAATCAAAAGTGGCCGAGGAGTCTTG 1380  
DB 1321 GCAGCCAGAACAGATATTTCAAGTGAAGCTGACAAATCAAAAGTGGCCGAGGAGTCTTG 1380  
OY 1381 GTGCTGTGAGGACAAAGGCTCAATCCAAATTTGTCCCTTAACCTGATTTGGCGATCC 1440  
DB 1381 GTGCTGTGAGGACAAAGGCTCAATCCAAATTTGTCCCTTAACCTGATTTGGCGATCC 1440  
OY 1441 AACACTCATGCCAGACATTTGATGCCAAGAGCTTGACCCAAAGTCCAGGACATGATGATCT 1500  
DB 1441 AACACTCATGCCAGACATTTGATGCCAAGAGCTTGACCCAAAGTCCAGGACATGATGATCT 1500  
OY 1501 CTGGGAGCTTAAATTTGAGAAAGAAAGCTCTCAGCCCAAGGTGAGAGGTGCGTCCATA 1560  
DB 1501 CTGGGAGCTTAAATTTGAGAAAGAAAGCTCTCAGCCCAAGGTGAGAGGTGCGTCCATA 1560  
OY 1561 GCCTCCATTGATGAGAGGCGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTT 1620  
DB 1561 GCCTCCATTGATGAGAGGCGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTT 1620  
OY 1621 TTTTTCATTAAGTTTGAAGAACCCGAATTC 1654  
DB 1621 TTTTTCATTAAGTTTGAAGAACCCGAATTC 1654

RESULT 2				
Locus	ARI41889			
Definition	Sequence 1 from patent US 6147194.	1642 bp	DNA	linear PAT: 08-AUG-2001
Version	ARI41889			
Keywords	ARI41889.1 GI:15101405			
Source	.			
Organism	Unknown:			
Reference	Unclassified. 1 (bases 1 to 1642)			
Authors	Collart,F.R. and Huberman,E.			
Title	Eukaryotic IMPDH polynucleotide and antibody compositions Patent: US 6147194-A I 14-Nov-2000;			
Journal	Location/Qualifiers 1..1642			
Features	source	/organism="unknown"		
Base Count	411 a      416 c      449 g      366 t			
Origin				

Query Match	98.98	Score 1635.6	DB 6	Length 1642
Best Local Similarity	99.88	Pred. 0.0		
Matches 1638	Conservative	0	Mismatches 4	Indels 0
				Gaps 0

OY	7	GGGGGCTCTCGGAGACACCGGGGGGTGCTCTGTGTGGGCATCGCC	ACTACCTGATTA	66
Db	1	GGCGGCTCTCGAGACACCGGGGGGTGCTCTGTGTGGGCATCGCC	ACTACCTGATTA	60
OY	67	GTGGGGGACGCTCTACGTGGCCAGACAGCGACATCACAGCACACAG	CTTCAACTGCG	129
Db	61	GTGGGGGACGCTCTACGTGGCCAGACAGCGACATCACAGCACACAG	CTTCAACTGCG	120
OY	127	GAGAGGCGCTACCTACACATGACTTCTCATCTTCCTGGGTACATCG	ACTTCACTCAG	186
Db	121	GAGAGGCGCTACCTACACATGACTTCTCATCTTCCTGGGTACATCG	ACTTCACTCAG	180
OY	187	ACCAAGTGGACCTGACTTCTGCTTGACCAAGAAATACATCTTAAAG	CCCCACATGGTTT	246
Db	181	ACCAAGTGGACCTGACTTCTGCTTGACCAAGAAATACATCTTAAAG	CCCCACATGGTTT	240
OY	247	CCTGTCCCATGAGACACAGTCACAGAGGTGGATGAGGCATAGCAATG	CCGTTACAGCG	306
Db	241	CCTGTCCCATGAGACACAGTCACAGAGGTGGATGAGGCATAGCAATG	CCGTTACAGCG	300
OY	307	GTAATTGGCTTCATCCACCACAACTGTACACCTGTAATTCAGCGCCAA	TAAATTGCGAAG	366
Db	301	GTAATTGGCTTCATCCACCACAACTGTACACCTGTAATTCAGCGCCAA	TAAATTGCGAAG	360
OY	367	TGAAGAAATATGACACAGGATTCATACACACCTGTGGTCTCAGCC	CCAAGATTCGG	428
Db	361	TGAAGAAATATGACACAGGATTCATACACACCTGTGGTCTCAGCC	CCAAGATTCGG	420
OY	427	TGCGGGATGTTTTGAGGCCAAGGCCCGGATGGTTCTCGCGTATCG	TAATCACAACA	486
Db	421	TGCGGGATGTTTTGAGGCCAAGGCCCGGATGGTTCTCGCGTATCG	TAATCACAACA	480
OY	487	CAGGCCGATGCGGAGCGCCGCTGTGGGGCATCATCTCTCCAGGGAC	TTGATTTTCTCA	546
Db	481	CAGGCCGATGCGGAGCGCCGCTGTGGGGCATCATCTCTCCAGGGAC	TTGATTTTCTCA	540
OY	547	AAGAGAGAGAACATGACTGTTTTGTGAAAGATTAATGACAAAGAGG	AGAGCTTGGTG	606
Db	541	AAGAGAGAGAACATGACTGTTTTGTGAAAGATTAATGACAAAGAGG	AGAGCTTGGTG	600
OY	607	TAGCCCCCGCAGCATCACACCTGAGGAGGCAATGAATTTCTCAGG	GCAGCAAGAAG	666
Db	601	TAGCCCCCGCAGCATCACACCTGAGGAGGCAATGAATTTCTCAGG	GCAGCAAGAAG	660
OY	667	GAAGTTGGCCATGTAAATGAACATGATGAGCTTGGGGCATCATTTG	CCCGGACAGAC	726
Db	661	GAAGTTGGCCATGTAAATGAACATGATGAGCTTGGGGCATCATTTG	CCCGGACAGAC	720
OY	727	TGAAGAAATGCGGACTCCCACTAGCTCCCAAGATGCCAAGAAATG	AGCTCTGTGTG	786

[illegible]

Unclassified.  
1 (bases 1 to 1642)  
AUTHORS Collart,F.R. and Huberman,E.  
TITLE METHOD AND MATERIALS RELATING TO IMPDH AND GMP PRODUCTION  
JOURNAL Patent: WO 9001545-A 1-22-FEB-1990;  
FEATURES Location/Qualifiers  
source 1..1642  
/organism="unknown"  
BASE COUNT 411 a 416 c 449 g 366 t  
ORIGIN  
Query Match 98.9%; Score 1635.6; DB 6; Length 1642;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 GGGCGGCTCTCGGAGACACAGCGCGGTGTCGTTGGCCATCGGCGCATCTGATTA 66  
DB 1 GGGCGGCTCTCGGAGACACAGCGCGGTGTCGTTGGCCATCGGCGCATCTGATTA 60  
QY 67 GTGGGGGCGACGTCTACGTGCGCAGACGACGACGACGACGACGACGACGACGACG 126  
DB 61 GTGGGGGCGACGTCTACGTGCGCAGACGACGACGACGACGACGACGACGACGACG 120  
QY 127 GAGAGCGGCTCACCTACATGACTTTCTATCTCTCCCTGGTACATCGACTTCAGTGCAG 186  
DB 121 GAGAGCGGCTCACCTACATGACTTTCTATCTCTCCCTGGTACATCGACTTCAGTGCAG 180  
QY 187 ACCAGGTGACCTGACTTGTGCTGCTGACCAAGAAATCACTCTTAAGACCCCACTGGTTT 246  
DB 181 ACCAGGTGACCTGACTTGTGCTGCTGACCAAGAAATCACTCTTAAGACCCCACTGGTTT 240  
QY 247 CCTCTCCCATGACACAGTACACAGGCTGGATGGCCATAGCAATGGCGTTACAGCGG 306  
DB 241 CCTCTCCCATGACACAGTACACAGGCTGGATGGCCATAGCAATGGCGTTACAGCGG 300  
QY 307 GTATTGGCTTCATCCACCACTGTAACCTGAATTCAGGGCCATGAAAGTTGGGAAAG 366  
DB 301 GTATTGGCTTCATCCACCACTGTAACCTGAATTCAGGGCCATGAAAGTTGGGAAAG 360  
QY 367 TGAAGAATATGAACAGGAGTTTCATACAGACCCCTGTGCTCAGACCCCAAGATCGCG 426  
DB 361 TGAAGAATATGAACAGGAGTTTCATACAGACCCCTGTGCTCAGACCCCAAGATCGCG 420  
QY 427 TCGGGATGTTTTTGAAGGCCAAGCGCCGCGATGGTTTCTGCGGTATCCCATACAGACA 486  
DB 421 TCGGGATGTTTTTGAAGGCCAAGCGCCGCGATGGTTTCTGCGGTATCCCATACAGACA 480  
QY 487 CAGGCGGATGGGGAGCCGCTTGGTGGGCTCATCTCCCTCAGGGACATTTGATTTCTCA 546  
DB 481 CAGGCGGATGGGGAGCCGCTTGGTGGGCTCATCTCCCTCAGGGACATTTGATTTCTCA 540  
QY 547 AAGAGGAGAACTGACTGTTTTCTTGGAAAGAGATTAATGACAAGAGGGAAGACTTGGTGG 606  
DB 541 AAGAGGAGAACTGACTGTTTTCTTGGAAAGAGATTAATGACAAGAGGGAAGACTTGGTGG 600  
QY 607 TAGCCCCCGCAGACATCACACTGAAGAGGCAATGAATTTCTGCAGCGCAGCAAGAGG 666  
DB 601 TAGCCCCCGCAGACATCACACTGAAGAGGCAATGAATTTCTGCAGCGCAGCAAGAGG 660  
QY 667 GAAGATTTGCCATTTGAATGAAGATGATGAGCTGTGGGCATCATTTGCCCGGACAGACC 726  
DB 661 GAAGATTTGCCATTTGAATGAAGATGATGAGCTGTGGGCATCATTTGCCCGGACAGACC 720  
QY 727 TGAAGAAGATCGGAGCTACTACCTTCACAAAGATGCCAAGAAACAGCTGCTGTGTG 786  
DB 721 TGAAGAAGATCGGAGCTACTACCTTCACAAAGATGCCAAGAAACAGCTGCTGTGTG 780  
QY 787 GGGCAGGCATTTGGCACTCATGAGATGACAAGATATAGCTGTGACTTCTGCCCCAGGCTG 846  
DB 781 GGGCAGGCATTTGGCACTCATGAGATGACAAGATATAGCTGTGACTTCTGCCCCAGGCTG 840  
QY 847 GTGTGATGTAGTGTGTTTTGGACTCTTCCACAGGAAATTCATCTTCCAGATCAATATGA 906  
DB 841 GTGTGATGTAGTGTGTTTTGGACTCTTCCACAGGAAATTCATCTTCCAGATCAATATGA 900

DB 841 GTGTGATGTAGTGTGTTTTGGACTCTTCCACAGGAAATTCATCTTCCAGATCAATATGA 900  
QY 907 TCAAGTACATCAAGACAAATACCTTAATTCACAAAGTCAATTTGAGAGCAATGTGCTACTG 966  
DB 901 TCAAGTACATCAAGACAAATACCTTAATTCACAAAGTCAATTTGAGAGCAATGTGCTACTG 960  
QY 967 CTGCCAGGCCCAAGAACTCAATGTATGCAAGGTGAGAGCCCTGGGGGTGGGCAATGGGAA 1026  
DB 961 CTGCCAGGCCCAAGAACTCAATGTATGCAAGGTGAGAGCCCTGGGGGTGGGCAATGGGAA 1020  
QY 1027 GTGGCTCATCTGCAATTAATCCAGAAAGTGTGGCTGTGGCGGGCCCAAGCAACAGCAG 1086  
DB 1021 GTGGCTCATCTGCAATTAATCCAGAAAGTGTGGCTGTGGGGGGGCCCAAGCAACAGCAG 1080  
QY 1087 TGTACAAAGTGTATGATGATGACGCGCCCTTTGCTGCTCCGCTCAATTTCTGATGAGGAA 1146  
DB 1081 TGTACAAAGTGTATGATGATGACGCGCCCTTTGCTGCTCCGCTCAATTTCTGATGAGGAA 1140  
QY 1147 TCCAAATATGTGGGTGATATTTGGAAGGCTTGGCCCTTGGGGCCCTCCACAGTCAATGATGG 1206  
DB 1141 TCCAAATATGTGGGTGATATTTGGAAGGCTTGGCCCTTGGGGCCCTCCACAGTCAATGATGG 1200  
QY 1207 GCTCTCTCTGCTGCTGCCACCACTGAGGCCCTGTGTAATCTTCTTTTCCGATGGGATCC 1266  
DB 1201 GCTCTCTCTGCTGCTGCCACCACTGAGGCCCTGTGTAATCTTCTTTTCCGATGGGATCC 1260  
QY 1267 GGCTAAAGAAATATGCGCGTATAGGTTCTCTGATGATGATGACAAAGACCTCAGACGCC 1326  
DB 1261 GGCTAAAGAAATATGCGCGTATAGGTTCTCTGATGATGATGACAAAGACCTCAGACGCC 1320  
QY 1327 AGACAGATATTTCAAGTGAAGCTGACAAATCAAAATGAGGCCAGAGAGTGTGCTGCTG 1386  
DB 1321 AGACAGATATTTCAAGTGAAGCTGACAAATCAAAATGAGGCCAGAGAGTGTGCTGCTG 1380  
QY 1387 TGCAGACCAAAAGGTCATATCACAATTTTGTCTTACCTGATTTGCTGGCATCAACT 1446  
DB 1381 TGCAGACCAAAAGGTCATATCACAATTTTGTCTTACCTGATTTGCTGGCATCAACT 1440  
QY 1447 CATGCCAGACATTTGTCGCAAGAGCTTTACCAAGTCCGAGACCATGATGTACTGTGGG 1506  
DB 1441 CATGCCAGACATTTGTCGCAAGAGCTTTACCAAGTCCGAGACCATGATGTACTGTGGG 1500  
QY 1507 AGCTTAAGTTTGAAGAGAGACGTCCTCAGCCAGGTGGAAGGTGGCGTCCATATAGCTTC 1566  
DB 1501 AGCTTAAGTTTGAAGAGAGACGTCCTCAGCCAGGTGGAAGGTGGCGTCCATATAGCTTC 1560  
QY 1567 ATTGCTATGAGAAAGCGGCTTTCTGAAAGGGAATCCAGACACCTCTCGGTTTTTTTTT 1626  
DB 1561 ATTGCTATGAGAAAGCGGCTTTCTGAAAGGGAATCCAGACACCTCTCGGTTTTTTTTT 1620  
QY 1627 CAATTAAGTTTGAAGAGACC 1648  
DB 1621 CAATTAAGTTTGAAGAGACC 1642

RESULT 4  
LOCUS BC006124 1663 bp mRNA linear PRI 12-JUL-2001  
DEFINITION Homo sapiens, IMP (inosine monophosphate) dehydrogenase 2, clone  
ACCESSION BC006124  
VERSION BC006124.1 GI:13543972  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1663)  
AUTHORS Strausberg,R.  
TITLE Direct Submision  
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK  
COMMENT

USA  
NIH-MGC Project URL: <http://imgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: AFCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consor  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@cgsc.bc.ca](mailto:info@cgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaro  
Susanna Chan, Readman Chiu, Chris Fjell, Erin  
Leticia Hsiao, Martin Krzywinski, Reta Kutsch  
Sen Lee, Victor Ling, Carrie Mathewson, Candic  
Ness, Pawan Pandoh, Anna-Jissa Prabh, Parvane  
Schein, Duane Smalrus, Michael Smith, Lorraine  
Michael Thorne, Miranada Tsai, Natasja van den  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IFAL Plate: 18 Row: f Column: 15  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency OK  
analysis, Genomescan gene prediction.

FEATURES  
source

Location/Qualifiers  
1..1663  
/organism="Homo sapiens"  
/db\_xref="LocusID:3615"  
/db\_xref="taxon:9606"  
/clone="MGC:13021 IMAGE:3545693"  
/rissue\_type="Colon, adenocarcinoma"  
/clone\_lib="NIH\_MGC\_15"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
41..1585  
/codon\_start=1  
/product="IMP (inosine monophosphate) dehydrogenase 2"  
/protein\_id="AAH06124.1"  
/db\_xref="GI:13543973"  
/translation="MADYLISGTSYVPDDLI7AOL7NCSDGLTYNDFLILPGYIDF  
TADVDLTSALTKITLTKPLVSSPMDDVETAGMA7IALTGIDIFIHNCPEFOAN  
EVRYKKEOGFIDPVLSPKRDYRDFEAKARHG7GIPITDGRKSLVGIIS  
RDIFLKEEHDFLEELMTKREDLVVAPAGITTLKEA7HEILORSKKGSLPIVNEDEL  
VAIIARMDLKKRNDYPLASKDKAKOLLCGA7IGHEDDKYRLDLQAQGVVYVLDSS  
OCNSIFQINM7IKYTKDKYPNLOVIGSNVY7AAQAKNLDAGVDALRVMSGSTCTTQ  
EVLACGRPOATAV7IKVSEYARFVGV7IADG10VNG7IAALALGASTVMGSL7AA  
TTEARGEVF7SDGIRLKKYRGMGSLDAMD7HLSQNR7FSEADK7IKVQGVSG7AVQDK  
GSIHKFV7YLIAG7IOHSCOD7IGAKSL7QVRAM7YSGE7KFEKRTSS7AOVEG7VHSLHS  
YEKRLF"

CDS

BASE COUNT 436 a 412 c 448 g 367 t  
ORIGIN

Query Match 98.1%; Score 1622.6; DB 9; Length 1663;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1628; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 GCGGCTCTGGAGACACGGCGGCTCTCTGTTGGCCATGGCGACACCTGATTAGT 68  
DB 2 GCGGCTCTGGAGACACGGCGGCTCTCTGTTGGCCATGGCGACACCTGATTAGT 61  
QY 69 GGGGACACGCTCTAGTGCAGAGACGAGACTACAGCAGCAGCAGCTCTCAACTGCGGA 128  
DB 62 GGGGACACGCTCTAGTGCAGAGACGAGACTACAGCAGCAGCAGCTCTCAACTGCGGA 121  
QY 129 GAGGCGCTCTAGTACATGACTTTCTCATCTCCCTGGGTACATGAGCTTCACTGACAC 188  
DB 122 GAGGCGCTCTAGTACATGACTTTCTCATCTCCCTGGGTACATGAGCTTCACTGACAC 181  
QY 189 CAGGTGACCTACTTCTGCTGACCAAGAAATCACTCTTAAGACCCACATGGTTTCC 248  
DB 182 CAGGTGACCTACTTCTGCTGACCAAGAAATCACTCTTAAGACCCACATGGTTTCC 241

QY 249 TCTCCCATGAGACACAGTCTACAGAGCGTGGGATGGCCATTAGCAATGGCGTTACAGCGGT 308  
DB 242 TCTCCCATGAGACACAGTCTACAGAGCGTGGGATGGCCATTAGCAATGGCGTTACAGCGGT 301  
QY 309 ATTGGCTTCATCCACCACACACTGTACACTGTAATTCAGAGCCAA7GAAGTTGGAAAGTG 368  
DB 302 ATTGGCTTCATCCACCACACACTGTACACTGTAATTCAGAGCCAA7GAAGTTGGAAAGTG 361  
QY 369 AAGAAATATGAACAGGAGATTCTACACAGACCTGTGTCTTCAGCCCCAAGATGCGGTG 428  
DB 362 AAGAAATATGAACAGGAGATTCTACACAGACCTGTGTCTTCAGCCCCAAGATGCGGTG 421  
QY 429 CGGATGTTTGGAGCCCAAGCCCGGCGATGGTTTCTGCGGTATCCCAATCAGACACA 488  
DB 422 CGGATGTTTGGAGCCCAAGCCCGGCGATGGTTTCTGCGGTATCCCAATCAGACACA 481  
QY 489 GCGCGATGGG3AGCCGCTTGTGGGCATCTCTCCAGGACAT7GATTTTCTCAA 548  
DB 482 GCGCGATGGG3AGCCGCTTGTGGGCATCTCTCCAGGACAT7GATTTTCTCAA 541  
QY 549 GAGGAGAACATGACTGTTCTTGGAAAGATTAATGACAAAGGGAACACTTGGTGTGA 608  
DB 542 GAGGAGAACATGACTGTTCTTGGAAAGATTAATGACAAAGGGAACACTTGGTGTGA 601  
QY 609 GCGCCCGCAGCATCACACTGAAGAGGCAAT7GAAATTC7CAGCGCAGCAAGAGGA 668  
DB 602 GCGCCCGCAGCATCACACTGAAGAGGCAAT7GAAATTC7CAGCGCAGCAAGAGGA 661  
QY 669 AAGTTGCCATTTGTAATGAATGATGAGCTTGTGGCCATAT7TCCCGGACACACTTG 728  
DB 662 AAGTTGCCATTTGTAATGAATGATGAGCTTGTGGCCATAT7TCCCGGACACACTTG 721  
QY 729 AAGAAGAATCGGAGACTACCCACTAGCCTTCCAAGATGCCCAAGAACAGCTGCTGTGGG 788  
DB 722 AAGAAGAATCGGAGACTACCCACTAGCCTTCCAAGATGCCCAAGAACAGCTGCTGTGGG 781  
QY 789 GCAGCATTTGGCATCATGAGATGACAAAGTATAGGCTGAGATTGCTCGCCAGCTGGT 848  
DB 782 GCAGCATTTGGCATCATGAGATGACAAAGTATAGGCTGAGATTGCTCGCCAGCTGGT 841  
QY 849 GTGATGTAGTGGTTTGGAGCTTCCAGGAGAAATTCATCTTCAGATCAATATGATC 908  
DB 842 GTGATGTAGTGGTTTGGAGCTTCCAGGAGAAATTCATCTTCAGATCAATATGATC 901  
QY 909 AAGTACATCAAAAGACAAATATCCCTAATCTCAAGTCATTTGGAGCAATGTGCTACTGCT 968  
DB 902 AAGTACATCAAAAGACAAATATCCCTAATCTCAAGTCATTTGGAGCAATGTGCTACTGCT 961  
QY 969 GCCCAGGCCAAGACCTCTTATGATGAGGTGAGATGCCCTCGGGGTGGGCAATGGGAAGT 1028  
DB 962 GCCCAGGCCAAGACCTCTTATGATGAGGTGAGATGCCCTCGGGGTGGGCAATGGGAAGT 1021  
QY 1029 GGCCTCATCTGATTAATCCAGAGAGTGTGCTGGCGGCGGCCCAAGCAACACAGCTG 1088  
DB 1022 GGCCTCATCTGATTAATCCAGAGAGTGTGCTGGCGGCGGCCCAAGCAACACAGCTG 1081  
QY 1089 TACAAGGTGTATGATATGACAGCGGCTTGTGTTCCGTCATTTGTCATGAGAGAAATC 1148  
DB 1082 TACAAGGTGTCAAGATATGACAGCGGCTTGTGTTCCGTCATTTGTCATGAGAGAAATC 1141  
QY 1149 CAAATATGCGGTCAATAT7TGGCAAAAGCCTTGGCCCTTGGCCCTCCACAGTCA7TGTGGGC 1208  
DB 1142 CAAATATGCGGTCAATAT7TGGCAAAAGCCTTGGCCCTTGGCCCTCCACAGTCA7TGTGGGC 1201  
QY 1209 TCTCTCTGTGCTGCCACCACTAGAGCCCTTGTGTAATCTTTCTTTCCGATGGATCCGG 1268  
DB 1202 TCTCTCTGTGCTGCCACCACTAGAGCCCTTGTGTAATCTTTCTTTCCGATGGATCCGG 1261  
QY 1269 CTAAGAAATAT7TCCGGGTATGGTTCTCTGATGCAATGGAACAAACACCTTACGACGCCAG 1328  
DB 1262 CTAAGAAATAT7TCCGGGTATGGTTCTCTGATGCAATGGAACAAACACCTTACGACGCCAG 1321  
QY 1329 AACAGATATTTAGTAGAAGCTGACAAATCAAGTGTGCCAGGAGTGTGTGCTCTGTG 1388

Dd	1322	AACACATATTTCCTAGTGAAGCTGACAAATCCAAAGTGCCCCAGGGAGGTGTCGTGCTCTGTG	1381
Oy	1389	CAGGACAAAAGGGTCAATCCACAAATTTGTCCCTTAACCTGATTGCCTGGCATCCAACACTCA	1448
Dd	1382	CAGGACAAAAGGGTCAANTCCACAATAATTTGTCCCTTAACCTGATTGCTGGCATCCAACACTCA	1441
Oy	1449	TGCCAGGACATTGGTGTCCCAAGACTTTGACCACCAAGTCCGAGCCTATATGACTCTGGGGAG	1508
Dd	1442	TGCCAGGACATTGGTGTCCCAAGACTTTGACCACCAAGTCCGAGCCTATATGACTCTGGGGAG	1501
Oy	1509	CTTAGTGTGAGAAGAGAAGCGTCCCTCAGGCCGAGTGGAGAAGGGTGCATGAGCTCCAT	1568
Dd	1502	CTTAGTGTGAGAAGAGAAGCGTCTTCAGGCCGAGTGGAGAAGGGTGCATGAGCTCCAT	1561
Oy	1569	TGCTATGAGAAAGCGGCTTTCTGAAAAGGATCCAGCACACCTCCTCGAGTTTTTTTTTCA	1628
Dd	1562	TGCTATGAGAAAGCGGCTTTCTGAAAAGGATCCAGCACACCTCCTCGAGTTTTTTTTTCA	1621
Oy	1629	ATAAAAAGTTTAGAANA	1645
Dd	1622	ATAAAAAGTTTAGAANA	1638

RESULT	5				
LOCUS	BC012840				
DEFINITION	BC012840	1648 bp	mRNA	linear	PRI 22-AUG-2001
ACCESSION	BC012840				
VERSION	BC012840.1	GI:15277479			
KEYWORDS	MGC.				
SOURCE	Homo sapiens.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	1 (bases 1 to 1648)				
TITLE	Straussberg, R.				
JOURNAL	Direct Submission				
	Submitted (20-AUG-2001) National Institutes of Health, Mammalian				

REMARK	COMMENT
NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>	
Contact: MGC help desk	
Email: <a href="mailto:cgapbs-rt@mail.nih.gov">cgapbs-rt@mail.nih.gov</a>	
Tissue Procurement: ATCC	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA library Arrayed by: The I.M.A.G.E. Consortium (HLNT)	
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center	
Center code: BCM-HGSC	
Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a>	
Contact: <a href="mailto:villallon@bcm.tmc.edu">villallon@bcm.tmc.edu</a>	
Villallon, D.R., Luna, R.A., Hele, S.M., Huiyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 12 Row: a Column: 6  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES	source	Location/Qualifiers
		1. 1648
		/organism="Homo sapiens"
		/db_xref="locusID:3615"
		/db_xref="taxon:9606"
		/clone="MGC:9332 IMAGE:3447994"
		/tissue_type="Placenta, Chorlocarcinoma"
		/clone_id="NH_MGC_10"
		/lab_host="DH10B"

[illegible]

Db	1	CTCGGAGACACGCGCGGTCTCTCTGTGTGGCCATGGCCGACTACCTGATTAGTGGGGGC	60
Qy	75	ACGTCCTACGTCGCGAGACGAGCTCACACACAGCAGCTTTCACACTGCGGAGAGCGC	134
Db	61	ACGTCCTACGTCGCGAGACGAGCTCACACACAGCAGCTTTCACACTGCGGAGAGCGC	120
Qy	135	CTCAGCTCAATGACTTTCTCATTTCTCCCTGGTACATCGACTTCAGTGCAGACCAAGT	194
Db	121	CTCAGCTCAATGACTTTCTCATTTCTCCCTGGTACATCGACTTCAGTGCAGACCAAGT	180
Qy	135	GACCTGACTTCGCTGCTGACCAAGAAATTCACCTCTTAAGACCCGACAGTTCCTGCCC	254
Db	161	GACCTGACTTCGCTGCTGACCAAGAAATTCACCTCTTAAGACCCGACAGTTCCTGCCC	240
Qy	255	ATGACACAGTCACAGAGCGCTGGGATGGCCATGACCAATGGCGCTTACAGGCGCTATTGGC	314
Db	241	ATGACACAGTCACAGAGCGCTGGGATGGCCATGACCAATGGCGCTTACAGGCGCTATTGGC	300
Qy	315	TTTCATCCACCACTGATACACCTGATATTCAGGCGCAATGAATTCGGAAGTGAAGAA	374
Db	301	TTTCATCCACCACTGATACACCTGATATTCAGGCGCAATGAATTCGGAAGTGAAGAA	360
Qy	375	TATGAAAGGGATTCATACAGCCCTGTGTGCTCAGCCCAAGAGATGGCTGCGGGAT	434
Db	361	TATGAAAGGGATTCATACAGCCCTGTGTGCTCAGCCCAAGAGATGGCTGCGGGAT	420
Qy	435	GTTTTTGAGGCGCAAGGCGCGCGCATGTTTTCGCGGTATCCCAATCACAGACACAGCGCG	494
Db	421	GTTTTTGAGGCGCAAGGCGCGCGCATGTTTTCGCGGTATCCCAATCACAGACACAGCGCG	480
Qy	495	ATGGGAGCGCGTGGTGGGATCATCTCTCCACAGGACATGATTCTCTCAAGAGAG	554
Db	481	ATGGGAGCGCGTGGTGGGATCATCTCTCCACAGGACATGATTCTCTCAAGAGAGAG	540
Qy	555	GAACTGACTGTTCTTCTGAGAGATATATGCAAAAGAGGAAAGCTGTGGTGGCCCC	614
Db	541	GAACTGACTGTTCTTCTGAGAGATATATGCAAAAGAGGAAAGCTGTGGTGGCCCC	600
Qy	615	GCGACATCACACTGAAGAGGCAATATGAATTCGACGCGCACAGAAAGGAAAGTTG	674
Db	601	GCGACATCACACTGAAGAGGCAATATGAATTCGACGCGCACAGAAAGGAAAGTTG	660
Qy	675	CCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGCGAGACCTGAAGAAG	734
Db	661	CCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGCGAGACCTGAAGAAG	720
Qy	735	AATGGGACTACCACTGACCTCCAAAGATGCCAAGAAACGCTGTGTGTGGGGCAGCC	794

Db	721	AATGGGACTA	ACCACCTAG	CCCTCA	AAAGTGG	CCAGAA	AAACACTG	CTG	GTGGG	CGAGCC	780	
QY	795	ATTGGCACT	CATGAG	GATGACA	GATATAG	CGCTG	GA	CTGCTG	CGCC	CG	854	
Db	781	ATTGGCACT	CAATG	AGATGACA	AGTATAG	CGCTG	GA	CTGCTG	CGCC	CG	840	
QY	855	GTAGGCTTT	GACTCT	CTCCAG	GGAATTC	ATCTTC	ACATCA	AT	TTGAT	CAAGTAC	914	
Db	841	GTAGGCTTT	GACTCT	CTCCAG	GGAATTC	ATCTTC	ACATCA	AT	TTGAT	CAAGTAC	900	
QY	915	ATCAAGACA	AAATAC	CCCTAAT	CTCCAG	TCATTTG	AGGCAAT	ATG	GGTC	CTGCTGCC	974	
Db	901	ATCAAGACA	AAATAC	CCCTAAT	CTCCAG	TCATTTG	AGGCAAT	ATG	GGTC	CTGCTGCC	960	
QY	975	GCCAAAGC	CAATGAT	GACAGG	TTGATG	CGCCGCGG	GTGGG	CA	TG	GAATGGCTCC	1034	
Db	961	GCCAAAGC	CAATGAT	GACAGG	TTGATG	CGCCGCGG	GTGGG	CA	TG	GAATGGCTCC	1020	
QY	1035	ATTCGATTA	TTCAGGA	AGTGC	TGGCTGT	GGGCGG	CCCAAGCA	AAACA	CACATG	TACAG	1094	
Db	1021	ATTCGATTA	TTCAGGA	AGTGC	TGGCTGT	GGGCGG	CCCAAGCA	AAACA	CACATG	TACAG	1080	
QY	1095	GTGATGAT	GTATGCA	CGGCGC	CTTGGTGT	CTCGGTCAT	TTGCTGAT	GGAA	TCCAAAT	1154		
Db	1081	GTGATGAT	GTATGCA	CGGCGC	CTTGGTGT	CTCGGTCAT	TTGCTGAT	GGAA	TCCAAAT	1140		
QY	1155	GTGGCTCAT	TATTTGGCA	AAACCTT	GGCCCTT	GGGCGCTC	CACAGT	CAATG	TTGGGCTCTC	1214		
Db	1141	GTGGCTCAT	TATTTGGCA	AAACCTT	GGCCCTT	GGGCGCTC	CACAGT	CAATG	TTGGGCTCTC	1200		
QY	1215	CTGGCTGC	CACTGAG	GGCCCTG	TGTGA	TACTTTT	CCGAT	GGG	TTCCGCTA	AAAG	1274	
Db	1201	CTGGCTGC	CACTGAG	GGCCCTG	TGTGA	TACTTTT	CCGAT	GGG	TTCCGCTA	AAAG	1260	
QY	1275	AAATATCG	CGGTAT	GGCTTCT	CTCATG	CCATG	GACAGC	AAACCTC	AGC	GGCAGA	CAACA	1334
Db	1261	AAATATCG	CGGTAT	GGCTTCT	CTCATG	CCATG	GACAGC	AAACCTC	AGC	GGCAGA	CAACA	1320
QY	1335	TATTTCACT	GAACCTG	ACAAAT	CAAAAGT	GGCCAGG	AGTGT	CGT	CTGTG	CAGAC	1394	
Db	1321	TATTTCACT	GAACCTG	ACAAAT	CAAAAGT	GGCCAGG	AGTGT	CGT	CTGTG	CAGAC	1380	
QY	1395	AAAGGTCAT	TCCAAAT	TTGTCC	CTTAC	CGTATG	GTGGCAT	CAAA	TACATCAT	GGCCAG	1454	
Db	1381	AAAGGTCAT	TCCAAAT	TTGTCC	CTTAC	CGTATG	GTGGCAT	CAAA	TACATCAT	GGCCAG	1440	
QY	1455	GACATTTG	TGCCAA	GAGCTT	GACCCAA	GTCCGAG	CCATGAT	ACTCT	GGGAGC	TTAAG	1514	
Db	1441	GACATTTG	TGCCAA	GAGCTT	GACCCAA	GTCCGAG	CCATGAT	ACTCT	GGGAGC	TTAAG	1500	
QY	1515	TTTGAAGA	GACAAG	CTTCTC	AGCCAG	GTGGA	AGTGGC	GTCCAT	ATG	CTAT	1574	
Db	1501	TTTGAAGA	GACAAG	CTTCTC	AGCCAG	GTGGA	AGTGGC	GTCCAT	ATG	CTAT	1560	
QY	1575	GAGAAGCG	CGCTTT	CTGAAA	AGGATCC	AGACAC	CTCTCG	GT	TTT	TTTCAAT	TAATA	1634
Db	1561	GAGAAGCG	CGCTTT	CTGAAA	AGGATCC	AGACAC	CTCTCG	GT	TTT	TTTCAAT	TAATA	1620
QY	1635	GTTTAGA	AAAGA	1645								
Db	1621	GTTTAGA	AAAGA	1631								

RESULT 6	BC015567	BC015567	1655 bp	mRNA	linear	PRI 09-OCT-2001
LOCUS	BC015567					
DEFINITION	BC015567					
	Homo sapiens, IMP (inosine monophosphate) dehydrogenase 2, clone					
	MC:20947 IMAGE:4576285, mRNA, complete cds.					
ACCESSION	BC015567					
VERSION	BC015567.1					
KEYWORDS	MC.					
SOURCE	Homo sapiens.					
ORGANISM	Homo sapiens					

REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK
1	(bases 1 to 1655)	Direct Submission		
1	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	Submitted (04-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
		NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
		Contact: MGC help desk		
		Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>		
		Tissue Procurement: Louis Staudt		
		cDNA Library Preparation: Rubin Laboratory		
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
		DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),		
		Galtherdsburg, Maryland,		
		Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a>		
		Contact: <a href="mailto:nisc_mgc@nih.gov">nisc_mgc@nih.gov</a>		
		Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McLooney, J.C., McDowell, E.E., Pearson, R., Snyder, B., Stanttip, S., Thomas, P.J., Tjongson, J.E., Touchman, J.W., Tsurguen, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.		
		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a>		
		Series: IRAL Plate: 30 Row: j Column: 8		
		This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.		
FEATURES	SOURCE	Location/Qualifiers		
		1..1655		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/db_xref="taxon:9606"		
		/clone="MGC:20947 IMAGE:4576285"		
		/tissue.type="Primary B-Cells from Tonsils"		
		/clone.lib="NIH-MGC_48"		
		/lab_host="DH10B-R"		
		/note="Vector: pOTB7"		
		41..1585		
		/codon_start=1		
		/product="Imp (inosine monophosphate) dehydrogenase 2"		
		/protein_id="AAH15567.1"		
		/db_xref="GI:15990412"		
		/translation="MADYILISGGSYVPDGLTAOGLFNGDGLTYNDELILPGXIDF TANOVYDLTSATRKTKITLPTLVSSPMDDTYEAKMALIMALTGICIFIHNCSTPEFOAN EVKVKVTEDEGFTIDPEVLVSPKDRNVPFAKARHFCGIPITDTGMSGRLVGITISS RIDLPFLKEEHEDFLEIMRKREDLVVAPAGITLKEANETLORSKRGKLPITVNEDEL VATILARDLKKRDYPLASDAKRLDGLGALITDHEBKYRLDLAAGVADVVLDS QGSITIDIMIKYIKDYKPYMLQYIGGVVVTAAQAKNIDAGVALLRYGMSGSCITTO EYACGCPQATAVYKAYKSEYARPGVPIADGGILSONGHIKALALATGYVMSGLAA TTPAPGPEFSDSIRLKYKRGSLMDMLHSLONNHYFEADKIKYAGCGSAPVADQ GSHIKRPVYLIAGIQHSCODIGAKSLTQVAMMYSCGLKREKRTSSAQVGGVSHLS YERLRP"		
CDS				
		BASE COUNT	430 a 410 c 450 g 365 t	
		ORIGIN		
		Query Match	97.6% Score 1614.6; DB 9; Length 1655;	
		Best Local Similarity	99.4%; Pred. No. 0;	
		Matches 1620; Conservative	0; Mismatches 9; Indels 0; Caps 0;	
Oy	17	CGGAGACACGGGGGCTCTGTTGGCCATGCGGACACTACTGATTTAGTGGGGGAC	76	
Db	10	CGGAGACACGGGGGCTCTGTTGGCCATGCGGACACTACTGATTTAGTGGGGGAC	69	
Oy	77	GTCCTACGTCGACAGACGAGCTACAGACACAGCAGCTCTTCAACTGGGAGACGGCT	136	
Oy	70	GTCCTACGTCGACAGACGAGCTACAGACACAGCAGCTCTTCAACTGGGAGACGGCT	129	

137 CACCTACATGACTTTCTCATTTCTCCCTGGGTACATCGACTTCACTGACAGACAGGTGA 196  
130 CACCTACATGACTTTCTCATTTCTCCCTGGGTACATCGACTTCACTGACAGACAGGTGA 189  
197 CCTGACTTCTGCTCTGACCAAGAAAATCACTTTAAAGACCCACTGGTTCTCTCCCAT 256  
190 CCTGACTTCTGCTCTGACCAAGAAAATCACTTTAAAGACCCACTGGTTCTCTCCCAT 249  
257 GGACACAGTACAGAGGGCTGGGATGGCCATAGCAATGGGGCTTACAGGGGATTTGGCTT 316  
250 GGACACAGTACAGAGGGCTGGGATGGCCATAGCAATGGGGCTTACAGGGGATTTGGCTT 309  
317 CATCCACCAACAACCTGATACCTGAAATCCAGGSCCAATGAAGTTGGAAAGTGAAGAATA 376  
310 CATCCACCAACAACCTGATACCTGAAATCCAGGSCCAATGAAGTTGGAAAGTGAAGAATA 369  
377 TGAACAGGATTCATACAGAGCCCTGTGGTCTCTCAAGCCCAAGATTCGGGTGGGATTT 436  
370 TGAACAGGATTCATACAGAGCCCTGTGGTCTCTCAAGCCCAAGATTCGGGTGGGATTT 429  
437 TTTTGAAGGCAAGGCCCGCATGTTTCTGCGGTATCCCAATCAGACAGACAGCCGAT 496  
430 TTTTGAAGGCAAGGCCCGCATGTTTCTGCGGTATCCCAATCAGACAGACAGCCGAT 489  
497 GGGGAGCCGCTTGTGGGCATCATCTCTCCAGGACATTTGATTTTTCAGAGAGAGA 556  
490 GGGGAGCCGCTTGTGGGCATCATCTCTCCAGGACATTTGATTTTTCAGAGAGAGA 549  
557 ACATGACGTTTCTGGAAGATATATGACAAAGAGGAGACTTGGTGGTGGCCCGCC 616  
550 ACATGACGTTTCTGGAAGATATATGACAAAGAGGAGACTTGGTGGTGGCCCGCC 609  
617 CAGCATCAGACTGAAGAGGCAATGAATTTGCGAGCCAGCAAGAGAGAAAGTTGCC 676  
610 AGGCATCAGACTGAAGAGGCAATGAATTTGCGAGCCAGCAAGAGAGAAAGTTGCC 669  
677 CATTTGTAATGAAGATGATGAGCTTGTGGCATCATTTGCCCGAGACAGACTTGAAGAGAA 736  
670 CATTTGTAATGAAGATGATGAGCTTGTGGCATCATTTGCCCGAGACAGACTTGAAGAGAA 729  
737 TCGGAGCTACCCAGTACGCTCCCAAGATGSCCAAGAAAGAGCTGTGGTGGGAGCCCAT 796  
730 TCGGAGCTACCCAGTACGCTCCCAAGATGSCCAAGAAAGAGCTGTGGTGGGAGCCCAT 789  
797 TCGGAGCTACGAGATGATGAGCTTGTGGCATCATTTGCCCGAGAGCTGTGGTGGATGT 856  
790 TCGGAGCTACGAGATGATGAGCTTGTGGCATCATTTGCCCGAGAGCTGTGGTGGATGT 849  
857 AGTGGTTTGGAGCTCTCCAGGAGAAATTCATCTTCCAGATCATATGATCAAGTACAT 916  
850 AGTGGTTTGGAGCTCTCCAGGAGAAATTCATCTTCCAGATCATATGATCAAGTACAT 909  
917 CAAAGCAATATCCCTAATCTCCAAGTCAATGGAGGAGATGGTACAGTCCGCCAGGC 976  
910 CAAAGCAATATCCCTAATCTCCAAGTCAATGGAGGAGATGGTACAGTCCGCCAGGC 969  
977 CAAAGCACTCATTTGATGAGTGTGGATCCCTGCGGGTGGGATGGGAAGTGGCTCAT 1036  
970 CAAAGCACTCATTTGATGAGTGTGGATCCCTGCGGGTGGGATGGGAAGTGGCTCAT 1029  
1037 CTGCATTTATCCAGGAAGTCTGGCTGTGGGCGGCCCAAGCAACAGAGTGTACAGGT 1096  
1030 CTGCATTTATCCAGGAAGTCTGGCTGTGGGCGGCCCAAGCAACAGAGTGTACAGGT 1089  
1097 GTATGAGTATGACAGGCGCTTTGGTGTCTCCGGTCAATTTGATGAGAGATCCAAATGT 1156  
1111 GTATGAGTATGACAGGCGCTTTGGTGTCTCCGGTCAATTTGATGAGAGATCCAAATGT 1149  
1090 GTACGAGTATGACAGGCGCTTTGGTGTCTCCGGTCAATTTGATGAGAGATCCAAATGT 1149  
1157 GGGTCATTTGCAAGAGCTTGGCCCTTGGGCGCTCCACAGTCAATGAGGCTCTCTCT 1216  
1150 GGGTCATTTGCAAGAGCTTGGCCCTTGGGCGCTCCACAGTCAATGATGAGGCTCTCTCT 1209

1217 GGCTGCCACCACTAGGCCCCCTGGTGAATACTTCTTTCCGATGGGATCCGGCTAAAGAA 1276  
1210 GGCTGCCACCACTAGGCCCCCTGGTGAATACTTCTTTCCGATGGGATCCGGCTAAAGAA 1269  
1277 ATATCGCGTATGGGTCTCTCGATGATGACAGACACCTCAGACGCCAGAACAGATA 1336  
1270 ATATCGCGTATGGGTCTCTCGATGATGACAGACACCTCAGACGCCAGAACAGATA 1329  
1337 TTTGAGTGAAGCTGACAAATCAAGTGGCCAGGAGTGTGTGGTGTGACAGACAA 1396  
1330 TTTGAGTGAAGCTGACAAATCAAGTGGCCAGGAGTGTGTGGTGTGACAGACAA 1389  
1397 AGGTCATATCCAAATTTGTCCCTTACTTATGTCCTGATGATCCAAACCTCATGCCAGAA 1456  
1390 AGGTCATATCCAAATTTGTCCCTTACTTATGTCCTGATGATGATCCAAACCTCATGCCAGAA 1449  
1457 CATTTGGTCCAAAGACTTGAAGCCAGTCCAGGACATGATGATGATGATGATGATGAT 1516  
1450 CATTTGGTCCAAAGACTTGAAGCCAGTCCAGGACATGATGATGATGATGATGATGAT 1509  
1517 TGAGAAGAGAAAGCTCTAGCCCAAGTGGAGTGGCTGCATAGCCCTCATTCGTATGA 1576  
1510 TGAGAAGAGAAAGCTCTAGCCCAAGTGGAGTGGCTGCATAGCCCTCATTCGTATGA 1569  
1577 GAAGCGGCTTTTCTGAAAAAGGATCCAGACACCTCTCGGTTTTTTTTCATTAATAAGT 1636  
1570 GAAGCGGCTTTTCTGAAAAAGGATCCAGACACCTCTCGGTTTTTTTTCATTAATAAGT 1629  
1637 TTAGAAGAA 1645  
1630 TTAGAAGAA 1638

RESULT 7  
AC122514  
LOCUS  
DEFINITION Mus musculus chromosome UNK clone R24-455L10, WORKING DRAFT  
AC122514  
AC122514 .2 GI:22004610  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 225727)  
TITLE The sequence of Mus musculus clone unpublished  
JOURNAL 2 (bases 1 to 225727)  
REFERENCE  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE  
AUTHORS 3 (bases 1 to 225727)  
TITLE Direct Submission  
JOURNAL Submitted (30-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
COMMENT  
On Jul 30, 2002 this sequence version replaced gi:21105978.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
Project Information  
Center project name: M\_BB0455L10  
----- Summary Statistics -----  
Sequencing vector: M13: 0%  
Sequencing method: plasmid: 100%  
Chemistry: Dye-primer ET: 0% of reads



```
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990319
Consensus quality: 224286 bases at least Q40
Consensus quality: 224589 bases at least Q30
Consensus quality: 224737 bases at least Q20
Insert size: 174000: agarose-fp
Insert size: 225627: sum-of-ctrls
Quality coverage: 15.05 in Q20 bases: agarose-fp
Quality coverage: 11.65 in Q20 bases: sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It
* consists of 2 contigs. The true order of the
* is not known and their order in this sequence
* is arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished
* as soon as it is available and the accession
* be preserved.
* 1 53314: contig of 53314 bp in length
* 53315 53414: gap of unknown length
* 53415 225727: contig of 172313 bp in length.
FEATURES
source
1. 225727
location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP24-455L10"
1. 53314
misc_feature
/feature="assembly_name:Contig10"
53415..225727
misc_feature
/feature="assembly_name:Contig11"
53415..225727
BASE COUNT 70387 a 42350 c 42349 g 70510 t 131 others
ORIGIN
Query Match 80.2% Score 1327; DB 2; Length 225727;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1444; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
OY 7 GGGCGGCTCTCGAGACACGCGCGCTGTCCTGTTGGCCATGGCCG CTACCTGATTA 66
DB 109325 GGGCGGTTCTCGACACGCGCTGTCCTGTTGGCCATGGCCG CTACCTGATTA 109384
OY 67 GTGGGGGCGCTCTCTACCTGTCAGACAGAGGATCAGACAGCAGGCTTCACTGCG 126
DB 109385 GGGAGGCGCCTCTCTACCTGTCAGACAGAGGATCAGACAGCAGGCTTCACTGCG 109444
OY 127 GAGACGCGCTCTCTACCTGTCAGACAGAGGATCAGACAGGATCAGACAGGCTTCACTGCG 186
DB 109445 GGGAGGCGCCTCTCTACCTGTCAGACAGAGGATCAGACAGGATCAGACAGGCTTCACTGCG 109504
OY 187 ACCAGGTGACCTGCTCTCTGACCAAGAAATCAGCTTTTAA 246
DB 109505 ATCAGGTGACCTGCTCTCTGACCAAGAAATCAGCTTTTAA 109564
OY 247 CCTCTCCATGACACAGTCTCTGACCAAGAAATCAGCTTTTAA 306
DB 109565 CCTCTCCATGACACAGTCTCTGACCAAGAAATCAGCTTTTAA 109624
OY 307 GTATGGCTTATCCACACCACTGATGATTCAGGCGCATG 366
DB 109625 GTATGGCTTATCCACACCACTGATGATTCAGGCGCATG 109684
OY 367 TGAACAAATATGACAGGATTCATCAGACAGCCTGTCCTCAGC 426
DB 109685 TGAACAAATATGACAGGATTCATCAGACAGCCTGTCCTCAGC 109744
OY 427 TGGGGATGTTTATGAGCCCAAGCGCGCATGTTTGGCGTATCC 486
DB 109745 TGGGGATGTTTATGAGCCCAAGCGCGCATGTTTGGCGTATCC 109804
OY 487 CAGGCGGATGAGGAGCGCTTGTGGGATCATCTCCCTCAGGAGACA 546
DB 109805 CAGGCGGATGAGGAGCGCTTGTGGGATCATCTCCCTCAGGAGACA 109864
```

```
OY 547 AAGAGAGACATATCATCTGTTTCTTGAAGAGATATGACAAAGGAGAGACTTGTGG 606
DB 109865 AAGAGAGAGATATCATCGGTTCTTGAAGAGATATGACAAAGGAGAGACTTGTGG 109924
OY 607 TGGCCCCCGGATGACATGACAGAGGAGGAAATGAAATTTCTGAGGCGCAGAGAGG 666
DB 109925 TGGCCCCCGGATGACATGACAGAGGAGGAAATGAAATTTCTGAGGCGCAGAGAGG 109984
OY 667 GAAAGTTGCCCATTTGTAATGAGATGATGAGCTTGTGGCCATCTTCCCGGACAGC 726
DB 109985 GAAAGTTGCCCATTTGTAATGAGATGATGAGCTTGTGGCCATCTTCCCGGACAGC 110044
OY 727 TGAACAAATGAGGAGATACACAGCTGATGAGCTTCCAAAGATGAGGAGAGAGGCTGCTG 786
DB 110045 TGAACAAATGAGGAGATACACAGCTGATGAGCTTCCAAAGATGAGGAGAGAGGCTGCTG 110104
OY 787 GGGCAGCATTTGACATGATGAGATGACAAAGATATGAGCTGTCGCGCCAGGCTG 846
DB 110105 GGGCAGCATTTGACATGATGAGATGACAAAGATATGAGCTGTCGCGCCAGGCTG 110164
OY 847 GTGTGATGATGAGTGTGTTTGGACCTCTCCAGGAGATTTCCATCTTCCAGATCATATGA 906
DB 110165 GTGTGATGATGAGTGTGTTTGGACCTCTCCAGGAGATTTCCATCTTCCAGATCATATGA 110224
OY 907 TCAAGTATCATCAAAAGCAATACCTTAATCTTCAAGTATGAGGAGCAATGTGTCACCTG 966
DB 110225 TCAAGTATCATCAAAAGCAATACCTTAATCTTCAAGTATGAGGAGCAATGTGTCACCTG 110284
OY 967 CTGCCAGGCGCAAGAACTCATATGATGAGAGTGTGATGCTTGGGGTGGGATGAGGAA 1026
DB 110285 CTGCCAGGCGCAAGAACTCATATGATGAGAGTGTGATGCTTGGGGTGGGATGAGGAA 110344
OY 1027 GTGCTCCATGCTGATTTATCCAGAGAGTCTGCTGTCGCGCGCCCAAGCAGCAGCAG 1086
DB 110345 GTGCTCCATGCTGATTTATCCAGAGAGTCTGCTGTCGCGCGCCCAAGCAGCAGCAG 110404
OY 1087 TGTACAAAGTGTATGATGATGAGAGGCGCTTGTGTTCCGCTATGCTATGAGAGAA 1146
DB 110405 TGTACAAAGTGTATGATGATGAGAGGCGCTTGTGTTCCGCTATGCTATGAGAGAA 110464
OY 1147 TCCAAATATGAGGATATTTGAGGAGGCTTGGGCGCTTCCAGAGATGATGAG 1206
DB 110465 TCCAAATATGAGGATATTTGAGGAGGCTTGGGCGCTTCCAGAGATGATGAG 110524
OY 1207 GCTCTCTCTGCTGACACACCTGAGGCGCTGTTGTTTCTTTTCCGATGAGGATTC 1266
DB 110525 GCTCTCTCTGCTGACACACCTGAGGCGCTGTTGTTTCTTTTCCGATGAGGATTC 110584
OY 1267 GAGTAAAGATATGAGGATGAGGCTTCTGATGAGGATGAGGATGAGGATGAGGATGAG 1326
DB 110585 GAGTAAAGATATGAGGATGAGGCTTCTGATGAGGATGAGGATGAGGATGAGGATGAG 110644
OY 1327 AGAACGATATTTTCAAGTCAAGCTGACAAATCAAAATGAGGCGGAGGATGCTGCTG 1386
DB 110645 AGAACGATATTTTCAAGTCAAGCTGACAAATCAAAATGAGGCGGAGGATGCTGCTG 110704
OY 1387 TGCAGAGCAAAAGGATGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1446
DB 110705 TGCAGAGCAAAAGGATGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 110764
OY 1447 CATGCGAGGAGATTTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1506
DB 110765 CATGCGAGGAGATTTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 110824
OY 1507 AGCTTAAGTTTATGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1566
DB 110825 AGCTTAAGTTTATGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 110884
OY 1567 ATTGCTATGAGAGGAGGCTTGTGAAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1626
DB 110885 ATTGCTATGAGAGGAGGCTTGTGAAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 110944
```

[illegible]

Db	266	CCTCACCCATGGACACTGTACACAGGGCTGGATTGGCCATCGCGATGGCGCTTTACAGAG	325
Qy	307	GTAATGGCTTCATCCACACCAACCTGTACACCTAAATTCACAGGCCAATGGAATTCGGAAAG	366
Db	326	GTAATTGGTTTCATCCACACCACTCCACACCTGAATTCACAGGCCAATGGAATTCGGAAAG	385
Qy	367	TGAGAAATATGAACAGGGATTCATCAACAGACCTGTGGTCTCAGGCCCAAGATCGCG	426
Db	386	TGAAGAAATATGAACAGGGATTCATCACTACCCCGTGGCTCTTACCCCAAGATCGCTG	445
Qy	427	TGCGGGATGTTTGGAGCCCAAGGCCCGGATGGTTTCGCGGTATCCCAATCCAGACA	486
Db	446	TACGGGATGTTTGGAGCCCAAAACCAGGCGATGGCTTCTGTGTATCCCATCCACAGTA	505
Qy	487	CAGGCCGATGGGAGGCGCGTGGTGGGACATCACTTCCTCCAGGACATTTGAATTTTCTCA	546
Db	506	CAGCGCGATGGGAGGCGATGGTGGGACATCACTTCCTCCAGGACATTTGAATTTTCTCA	565
Qy	547	AAGAGAGAAACATGACTGTTTCTTGGAGAAGATTAATGACAAAGAGGAAAGACTTGGTGG	606
Db	566	AGGAGAGAAACATGACCGGTTCTTGGAGAAGATTAATGACTAAGAGGAAAGATTTGGTGG	625
Qy	607	TAGCCCCCGGAGCATCAACACTGAGAGGCAATGAATTCGACGGCCAGCAAGAAG	666
Db	626	TCGCCCCCGGAGCATCACTGTAAGAAAGGCAATGAGATTTCGACGCAAGCTATAAAAG	685
Qy	667	GAAAGTTCCCATGTTAAATGAAGATATAGCTTGGGCAATCTTGGCCGGACAGACC	726
Db	686	GAAAGTTCCCATGTTGTAATGAATATATAGCTGAGCATCAATTTGCCCGGACAGACC	745
Qy	727	TGAAGAAAGATCGGGACTACCCACTAGACCTCCAAAGATGCCAAAGAAAGCTGTGTGTG	786
Db	746	TAAAGAAAGATCGATTACCTCCCGCTCCAAAGATGCCAAAGAAAGCTGTGTGTG	805
Qy	787	GGGCAAGCATTTGSGACATCATGAGATGACAAGTATAGCTGGACCTTGCGCCCAAGCTG	846
Db	806	GGGCAAGCATTTGGACATCATGAGATGACAAGTATAGCTGGACCTTGCGCCCTTGTG	865
Qy	847	GTCGATGATGATGGTGTGTTGGACCTCTCCCAAGGAAATTCACATTCGCAATTCATATGA	906
Db	866	GTGTGATGATGATGGTGTGTTGGACCTCTCCCAAGGAAATTCACATTCGCAATTCATATGA	925
Qy	907	TCAGATACATCAAGACAAATACCCATATCTCCAAAGTCATTTGAGGCAATGTGTCATG	966
Db	926	TCAAATACATCAAGAGAGATCCAGTCATCAAGGTCATTTGAGGCAATGTATGTCATG	985
Qy	967	CTGCCCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGCGCTCGCGGGTGGGCAATGGGAA	1022
Db	986	CTGCGCAAGGCCAAGAACCTCATATGATCAAGTGTAAATGCTTTTGCAAGTCGCGCATGGGAA	1041
Qy	1027	GTCGCTCATCTGTCATTTATCCAGGAAGTGTGTGCTGTGGGCGCGCCCAAGCAACAGCAG	1088
Db	1046	GTGCTTCATCTGTCATCAACCAGGAAGTGTGGCTGTGTGGCGCGCCCAAGCAACAGCAG	1107
Qy	1087	TGTACAAAGCTGTATGAGTATGCAAGGCGCTTGTGTCTCCGTCATTTGCTGATGAGGAA	1140
Db	1106	TGTACAAAGCTGTGAGTATGACCCGCTGTGTGTGTCTGTATTTGCTGATGAGGAA	1165
Qy	1147	TCCAAATGTGGTCATATTGGCCAAAGCCTTGGCCCTTGGGGCTTCACAGTCATGATGG	1206
Db	1166	TCCAAATGTGGTCATATTGGCCAAAGCCTTGGCTTGGGGCTTCACAGTCATGATGG	1222
Qy	1207	GCTCTCTCTCGGCGCCACCACTGAGGCCCTGTGTGAATACCTCTTTCCGATGGGATTC	1266
Db	1226	GCTCTCTCTCGGCGCCACCACTGAGGCCCTGTGTGAATACCTCTTTCTAGATGGGATTC	1285
Qy	1267	GAGTAAAGAAATATTCGCGGATATGGGTTCTCTGCATGCCATGCAAGCACTCCACAGCC	1328
Db	1286	GAGTAAAGAAATATCCGAGATATGGGTTCTTGTATGCCATGGCAACACTCTCAGCAGCC	1344
Qy	1327	AGAACGATTTTCACTGAGAGCTGACAAAATCAAAATGGCGCCAGGAGTGTCTGTGCTG	1386

Db 1346 AGAACGACTTCACTGAGCTGACAAATTCAGAGCTTCAGGGGCGAG 1405  
Qy 1387 TCGAGGACAAAGGCTCAATCCACAAATTTGTCCCTTACCTGATTTGCTG 1446  
Db 1406 TCGAGGACAAAGGCTCAATCCACAAATTTGTCCCTTACCTGATTTGCTG 1465  
Qy 1447 CATGCCAGACATTTGTCGCAAGAGCTTACCCAGCTCCGAGGACATGA 1506  
Db 1466 CCTGTAAGACATTTGTCGCAAGAGCTTACCCAGCTCCGAGGACATGA 1525  
Qy 1507 AGCTTAAGTTTGAAGAAGAACGCTTCTCAAGCCAGCTGGAAGGTGGCG 1566  
Db 1526 AGCTTAAGTTTGAAGAAGAACGCTTCTCTGCTCAAGTGGAGAGTGGCG 1585  
Qy 1567 ATTGATGATGAAGAGGCGTTTCTGAAAGAGATCCAGACACCTGCT 1626  
Db 1586 ATTGATGATGAAGAGGCGTTTCTGAAAGAGATCCAGATATATGCTT 1645  
Qy 1627 CATATAAGTTTGAAGAAGACCCGAA 1651  
Db 1646 AAAAGTTTGGAAAAAAGTGA 1670

RESULT 9  
LOCUS BC010314 1680 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, inosine 5'-phosphate dehydrogenase 2, clone MGC:6193  
ACCESSION BC010314  
VERSION BC010314.1 GI:16307530  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1680)  
Strausberg, R.  
Direct Submission  
Submitted (05-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amgdbcm.tmc.edu](mailto:amgdbcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

FEATURES  
Source Clon distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 5 Row: o Column: 15.  
Location/Qualifiers  
1..1680

/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="CZECR 11"  
/clone="MGC:6193 IMAGE:2646679"  
/tissue\_type="Mammary tumor metastasized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMTV."  
/clone\_lib="NCI\_CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"  
34..1578  
CDS

/codon\_start=1  
/product="Inosine 5'-phosphate dehydrogenase 2"  
/protein\_id="AAH0314.1"  
/db\_xref="GI:16307531"  
/db\_xref="LocusID:23918"  
/translation="MADYLISGTSVPDDGLTAQDLPNCGIGLTYNDFLLPGYIDF  
TADVDLSALTKTKTLTPRIVSSPMQVTEGMAIMATGCTGCTFHNCPTPEQAN  
EVRKRTKEQFIDPVLSPKDRVDFEAKARFGGCIPTDTGRKMSRLVGLISS  
RIDDLKEEHRFLIEIMTKREDLVAVATYLLKLANE1LDRSKKLIPIVNEDEL  
VAIIARTDLKKNRDPPLASDKAKKOLLGAAIGTHEDBKVRIDLALAGVDVLDSS  
QGNISFOIMIKKRYIKEYKPSLOVIGNVTAQAKNLLDAGVDALRVMGSSGICITQ  
EVLACGPOATAVRYEVARREYVADGDIQVNGHAKALALGASVMSGSLAA  
TEAPGEVFPDSGTRILKRYRCGMSLDAMDKHLISQNRPESEDKIKVAGVSGAODK  
GSIHKFPVYLINGIODIGAKSLTQVRAMRYSGELKFEKRTSAYOEGVSHLS  
YERKLP"

BASE COUNT 465 a 396 c 435 g 384 t  
ORIGIN

Query Match 79.9%; Score 1322.2; DB 10; Length 1680;  
Best local Similarity 88.2%; Pred. No. 0;  
Matches 1438; Conservative 0; Mismatches 193; Indels 0; Gaps 0;  
Qy 15 CTGGGAGACACGGCGGCTGCTCTGTTGGCATGGCCGACTTACTGATTTAGGGCG 74  
Db 1 CTGGGAGACACGGCGGCTGCTCTGTTGGCATGGCCGACTTACTGATTTAGGGCG 60  
Qy 75 AGCTCTAGGTCGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 134  
Db 61 ACCTCTAGGTCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 120  
Qy 135 CTGACCTACAGTACTTTCTCATTTCTCCCTGGGTACATGACTTCACTGACGACAGGTG 194  
Db 121 CTCACCTACAGTACTTTCTCATTTCTCCCTGGGTATATGACTTCACTGACGACAGGTG 180  
Qy 195 GACCTGACTTCTCTGACCAAGAAATACCTTTAAGACCCGCTGTTCTCTCC 254  
Db 181 GACTTACGCTCTCTTACTAAGAAATTTACTTAAAGCCCATTTGGTTCTTACC 240  
Qy 255 ATGACACAGCTCAGACAGCGTGGGATGCCATAGCAATGACGCGCTTACAGCGGATTTGCC 314  
Db 241 ATGACACAGCTCAGACAGCGTGGGATGCCATAGCAATGACGCGCTTACAGCGGATTTGCC 300  
Qy 315 TTGATCCACACAGCTGTACACTGATTCAGAGCCAAATGAAGTTGGAAAGGAAGAA 374  
Db 301 TTGATCCACACAGCTGTACACTGATTCAGAGCCAAATGAAGTTGGAAAGGAAGAA 360  
Qy 375 TATGAACAGGAGATTATCAGACAGACCTGTGCTCTGAGCCCAAGGATGCGTCCGGAT 434  
Db 361 TACGAACAGGAGATTATCAGACAGACCTGTGCTCTGAGCCCAAGGATGCGTCCGGAT 420  
Qy 435 GTTTTGGAGCCCAAGGCGCGGATGTTCTGCGGTATCCCAATCAGACAGCGCGG 494  
Db 421 GTTTTGGAGCCCAAGGCGCGGATGTTCTGCGGTATCCCAATCAGACAGCGCGG 480  
Qy 495 ATGGGAGCGCGTGTGGGATCATCTCTCCAGGACATTTGATTTTCAAGAGGAG 554  
Db 481 ATGGGAGCGCGTGTGGGATCATCTCTCCAGGACATTTGATTTTCAAGAGGAG 540  
Qy 555 GAACATGACTGTTTGGAAAGATATGACAAAGGAGGAGACTTGGTGTAGCCGCC 614  
Db 541 GACATGACCGGTTCTTGAAGAGATATGACTTGAAGGAGGAGATTGGTGTAGCCGCC 600  
Qy 615 CGCAGCATCACATGAAGAGGCAATTAATCTGACGCGCAGCAAGAGAGGAAAGTTG 674  
Db 601 GCGGCGCTACCTGTGAAGAGGCAATTAATGAATTTCTACAGCAAGTAAAAAGAGTTG 660  
Qy 675 CCATTTGAATGAAGATATGAGCTTGTGGCATATTTGCCGAGACAGCTGAAGAG 734  
Db 661 CCATTTGAATGAAGATATGAGCTTGTGGCATATTTGCCGAGACAGCTGAAGAG 720  
Qy 735 AATCGGACTACCACTTACCTTCAAGAGATGCCAAGAAACAGCTGCTGTGGCGGCGC 794  
Db 721 AATCGATTTACCCCTGCGCTCAAGAGATGCCAAGAAACAGCTGCTGTGGCGGCGC 780

795 ATTGGCACTCATGAGATGACAAAGTATAGCTGGAAGCTTCGCCAGCGCTGTGTGAT 854  
|||||  
781 ATTGGCACTCATGAGATGACAAAGTATAGCTGGAAGCTTCGCCAGCTGTGTGAT 840  
|||||  
855 GTAGTGGTTTTGGACTCTTCCAGGAAATTCATCTTCACATCAATATGATCAAGTAC 914  
|||||  
915 ATCAAGAACAATACCCTTAATCTCCAGATCATTTGGAGCAATGTGTCTACTGTGCCAG 974  
|||||  
901 ATCAAGGAACAATATCCCAAGTCTACAGTCATTTGGAGCAATGTATCTACTGTGGCAA 960  
|||||  
975 GCCAAGAACCTTATTTGATGAGTGGATGGCTTGGAGTGGGATGGGAATGGGCTCC 1034  
|||||  
961 GCCAAGAACCTTATTTGATGAGTGGATGGCTTGGAGTGGGATGGGAATGGGCTCC 1020  
|||||  
1035 ATCTGCATTATCCAGAACTGTGCTGGCTGGGCGGCCCAAGCAACAGCAGTGTACAA 1094  
|||||  
1021 ATCTGCATCACCAGAACTGTGCTGGCTGGGCGGCCCAAGCAACAGCAGTGTACAA 1080  
|||||  
1095 GTGTATGATATGACAGCGGCTTTGTGTCTCCGGTCAATTCGTATGGAGGAATCCAAAT 1154  
|||||  
1081 GTCTGTGAATATGCCCGTGTGTGTCTTCCGTATTTGCTGATGAGAGGAATCCAAAT 1140  
|||||  
1155 GTGGGTCATATTTGCCAAAGCTTTGGGCTTGGGCGCTCCACAGTCATGATGGGCTCTC 1214  
|||||  
1141 GTGGGTCATATTTGCCAAAGCTTTGGGCTTGGGCGCTCCACAGTCATGATGGGCTCTC 1200  
|||||  
1215 GTGGGTCACCACTGAGGCGCTTGGTGAATCTTCTTCCGATGGGATCCGGCTCAAG 1274  
|||||  
1201 GTGGGTCACCACTGAGGCGCTTGGTGAATCTTCTTCCGATGGGATCCGGCTCAAG 1260  
|||||  
1275 AATATCGGCGATATGGTCTCTCGATGGCATGAGCAACACCTCCAGCCAGCAACAGA 1334  
|||||  
1261 AATATCGGCGATATGGTCTCTCGATGGCATGAGCAACACCTCCAGCCAGCAACAGA 1320  
|||||  
1335 TATTTCAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTCTGTCTGTGACAGAC 1394  
|||||  
1321 TACTTCAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTCTGTCTGTGACAGAC 1380  
|||||  
1395 AAGAGCTCATTCACAAATTTTGTCTTACCTGATTTGTCTGTGACATCCACATCATGCCAG 1454  
|||||  
1381 AAGAGCTCATTCACAAATTTTGTCTTACCTGATTTGTCTGTGACATCCACATCATGCCAG 1440  
|||||  
1455 GACATTTGGTGAAGAGTGTGAAGTCCAGGCGCATGATGTACTCTGGGAGCTTAAG 1514  
|||||  
1441 GACATTTGGTGAAGAGTGTGAAGTCCAGGCGCATGATGTACTCTGGGAGCTTAAG 1500  
|||||  
1515 TTTTGAAGAAGAGAGTCTTCACAGCCAGGTGGAAGTGGGCTCCATGAGCTTCATTCGTAT 1574  
|||||  
1501 TTTTGAAGAAGAGAGTCTTCACAGCCAGGTGGAAGTGGGCTCCATGAGCTTCATTCGTAT 1560  
|||||  
1575 GAGAGCGGCTTTTCTGAAAAGGAGTCCAGACACTCTCGTGGTTTTTTTTCATTAATAA 1634  
|||||  
1561 GAGAGCGGCTTTTCTGAAAAGGAGTCCAGACACTCTCGTGGTTTTTTTTCATTAATAA 1620  
|||||  
1635 GTTTAGAAGA 1645  
|||||  
1621 GGGAAAAAAA 1631  
|||||

RESULT 10  
ARI41890 1620 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 2 from patent US 6147194.  
ACCESSION ARI41890  
VERSION ARI41890.1 GI:15101406  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1620)

AUTHORS Collart,F.R. and Huberman,E.  
TITLE Eukaryotic IMPDH polynucleotide and antibody compositions  
JOURNAL Patent: US 6147194-A 2 14-NOV-2000;  
FEATURES Location/Qualifiers  
source 1..1620  
BASE COUNT 423 a 369 c 421 g 407 t  
ORIGIN  
Query Match 79.9% Score 1321.8; DB 6; Length 1620;  
Best Local Similarity 88.8%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 177; Indels 4; Gaps 1;  
23 CACGGGCGGCTTCTCTGTTGGCCATGCGCCACTGATGATAGTGGGCGACGCTCTA 82  
|||||  
1 CACGGCTCGGTCTCTCTGTTGGCCATGCGCGACTGATGATAGTGGGCGACATCTTA 60  
|||||  
83 CGTCCAGACGACGAGCTCACAGCACAGCAGCTCTTCAACTGCGGAGACGCGCTCACCTA 142  
|||||  
61 CGTCCCGACGACGAGGCTCACAGCGCAGCCTTCAACTGCGGAGATGGCCTCACCTA 120  
|||||  
143 CATGACTTTCTCATTTCTCCCTGGGTACATCGACTTTCATGACAGACAGTGGACTGAC 202  
|||||  
121 CAACGATTTTCTCATTTCTCTGGGTATATGCACTTCATCGCGACAGTGGATTTGAC 180  
|||||  
203 TTCTGCTGACCAAGAAATCATCTTTAAGACCCCACTGTTTCTCTCCCATGAGAC 262  
|||||  
181 CTCTGCTTACTAAGAAATCATCCTTGAAAGACCACAGCTGTTTCTCTCATATGAGAC 240  
|||||  
263 AGTCACAGAGGCTGGGATGGCATAGCAATGGCGCTTACAGCGGATTTGGCTTATCCA 322  
|||||  
241 TGTACAGAGGCTGGAAATGGCCATTTGCAATGGCGCTTACAGAGGATTTGGCTTATCCA 300  
|||||  
323 CCACAACCTGACACCGTAATTCAGGCGCAATGAGTGGGAAGTGAAGAATATGAGACA 382  
|||||  
301 CCACAACCTGACACCTGTAATTCAGGCGCAATGAGTGGGAAGTGAAGAATATGAGACA 360  
|||||  
361 GGGATTCATTAACATGATCTGTAGTCTTACGCCCAAGATCGTGTGAGGAGTGTTTGA 420  
|||||  
443 GGGCAAGCGCGGATGTTTCTGCGGTATCCCATTCACAGACAGCGCGATGGGGAG 502  
|||||  
421 AGCCAAAGAGGAGATGAGTGTGTATCCCATTCACAGATACAGAGCGGATGGGGAG 480  
|||||  
503 CGGCTTGGTGGGATATCTCTCCAGGACATGTTTTCATAAGAGAGGAACATGA 562  
|||||  
481 TCGACTGTGGGATATCTTTCATAGGATATTTTTCATAGGAGAGAGATGA 540  
|||||  
563 CTGTTTCTTGAAGATATGACAAAGAGGGAAGCTTGTGTGAGCCCGGACGAT 622  
|||||  
541 CGTTTCTTGAAGATATGACAAAGAGGGAAGTTGTGTGAGCCCGGACGAT 600  
|||||  
623 CACACTGAAGAGGCAATATGAAATTTCTGACGCGCAGCAAGAGGAAATGTCCTAT 682  
|||||  
601 CACTCTGAAGAGGCAATATGAAATTTCTGACGCGCAGCAAGAGGAAATGTCCTAT 660  
|||||  
683 AATGAAGATGATGACTGTGGGCTCATGTTGGCGGACAGCTGAAGAAGATGGGA 742  
|||||  
661 GAATGAAGATGATGACTGTGGGCTCATGTTGGCGGACAGCTGAAGAAGATGGGA 720  
|||||  
743 CTACCCACTAGCTCCAAAGATGCCAAGAAACAGCTGTGTGGGCGACGATTTGGAC 802  
|||||  
721 TTACCCATTTGGCTTCCAAAGATGCCAAGAAACAGCTGATGTGGGCGACGATTTGGTAC 780  
|||||  
803 TCATGAGATGACAGTATAGCTGTGAGCTTCTGCGCCAGCGCTGTGTGATGTAGTGT 862  
|||||  
781 TCATGAGATGACAGTATAGCTGTGAGCTTCTGCGCCAGCGCTGTGTGATGTAGTGT 840  
|||||  
863 TTTGAGCTCTTCCAGGGAATTTCCATCTTCAGATCATATGATCAATACATCAAGA 922  
|||||  
841 TTTGAGCTCTTCCAGGGAATTCATTTTCCAAATCATATGATCAATACATCAAGA 900  
|||||

QY	923	CAAAATACCTAATCTCCAGTCATTGGAGGCAATGTGCTACTGCTGCACAGCCACAGAA	982
Db	901	GAATAACCCCAATCTCCAGTCATTGGAGGCAATGTAGTACCTGCTGCACAGCAAGAA	960
QY	983	CCTCATATGACGAGTGTGATGCCCTCGGGGTGGCATGGAGAGGTGCTCATCGCAT	1042
Db	961	CCTCATAGACGACGAGTGTGATGCTCTCGGAATTTGGCATGGGGTGTGATTCATCTGCAT	1020
QY	1043	TATCCAGGAATGTGGCCTGTGGGCGGCGCCCAAGCACAAGCATGTATTAAGGTGTATGA	1102
Db	1021	TACTAGGAAAGTGTGGCCTGTGGTGGCGCCCAACCAACAGCAGTGTATTAAGGTTTCTGA	1080
QY	1103	GTATGACGAGCCCTTTGGTGTCCGCTCATTTCTGTATGTGAGAAATCCAAATGTGGGTCA	1162
Db	1081	GTATGCTCGGCCCTTTGGTGTCTTCTGTATTCTGTATGTGAGAAATCCAAATGTGGGTCA	1140
QY	1163	TATTTGGAAAGCCTTTGGCCCTTGGGGCCTCCACAGTCATGATGGGCTGCTCTGGCTGC	1222
Db	1141	TATTTGGAAAGCCTTTGGCCTTGGAGCTTCTACATCATGATGGGCTGCTCTGGCTGC	1200
QY	1223	CACCACTAGGGCCCTGGTGAATACTCTTTTCCATGGGATCGGGCTTAAGAAATATCG	1282
Db	1201	CACCACTAGGGCCCTGGTGAATACTCTTCTCAGATGGGATCGGGCTTAAGAAATATCG	1260
QY	1283	CGGTATGGGTTCTCTGATGCCATGGACAAAGCACCCTCAGACCCAGATAGATATTTGAG	1342
Db	1261	TGGTATGGGTTCTCTGATGCCATGGACAAAGCACCCTCAGACCCAGATAGATATTTGAG	1320
QY	1343	TGAACCTGCAAAATCAAAAGTGCGCCAGGAGATGTCTGCTGTGCAAGCAAAAGGTTC	1402
Db	1321	TGAACCTGCAAAATCAAAAGTGCGCCAGGAGATGTCTGCTGTGCAAGCAAAAGGTTC	1380
QY	1403	AATCCACAATTTGTCCCTTACCTGATTGCTTGATCCACACCTCATCAGACGATTTGG	1462
Db	1381	TATCCACAAGTTGTCCCTTATTTGATTGCTTGATCCACACCTCATCAGACGATTTGG	1440
QY	1463	TGCCAAGACCTTGACCCCAAGTCCGAGCCATGTACTCTGGGGAGCTTAAGTTTGAGA	1522
Db	1441	TGCCAAGACCTTTAACCCCAAGTCAAGGACCATGTACTCTGGGGAGCTTAAGTTTGAGA	1500
QY	1523	GAGAACGTCCTCAGCCCAAGGTGAGAGTGGGCTCATAGCCTCAATTTGATGACAAGG	1582
Db	1501	GAGAACGTCCTCAGCTCAGGTGAGAGTGGGTTCACAGCCTTCATTTGATGACAAGG	1560
QY	1583	GCTTTTTCGAAAAGGATCTCAGACACCTCCCTCGGCTTTTTTTCGAATTAAGTTTGA	1642
Db	1561	GCTTTTTCGAAAAGGATCTCAGCTA---TATGCTTGAATTTTTCGAATTAAGTTTGA	1616
QY	1643	AGA 1645	
Db	1617	AAA 1619	
RESULT 11			
LOCUS	106572	1620 bp	DNA
DEFINITION	Sequence 3 from Patent WO 9001545.		Linear
ACCESSION	106572		
VERSION	106572.1	GI:589654	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1620)		
AUTHORS	Collart, F. R. and Huberman, E.		
TITLE	METHOD AND MATERIALS RELATING TO IMPDH AND GME PRODUCTION		
JOURNAL	Patent: WO 9001545-A 3 22-FEB-1990;		
FEATURES	Location/Qualifiers		
source	1..1620		
BASE COUNT	423 a 369 c 421 g 407 t		
ORIGIN	/organism="unknown"		

Query Match	79.9%	Score 1321.8	DB 6	Length 1620
Best Local Similarity	88.8%	Pred. No. 0		
Matches 1442	Conservative	0	Mismatches 177	Indels 4
Gaps				
QY	23	CACGGGGGGGGTGTCTGTGTGGCCATGGCCGACCTACCTGATATAGTGGGGGACGCTCCCA	82	
Db	1	CACGGGCTCGCTCTCTCTGTGGCCATGGGGGACCTACCTGATTAAGGGAGGACATCTTAA	60	
QY	83	CGTGGCAGACGACGAGACTCAGACGACAGAGCTCTTCAACTGCGGGAGAGGCGCTCACTCA	142	
Db	61	CGTGGCCGACGACGAGGCTCAGACGCGACAGCTCTTCAACTGCGGGGATGGCTCACTCA	120	
QY	143	CAATACCTTTTCATTTCTCCCTGGGTACATACGACTTACTGCGAGACCAGTGGACTGTAC	202	
Db	121	CAAGATTTTTCATTTCTCTCTGGGTATATCGACTTACTGCGGACCAAGTGGATTTGTAC	180	
QY	203	TTCTGCTGTGACCAAGAAATACACTTTAAGACCCCACTGGTTCCCTCCCATGTGACAC	262	
Db	181	CTCTGCTGTATCAAGAAAGATCACCTTGAAGACCCCACTGGTTCCCTCACTATATGACAC	240	
QY	263	AGTCACAAGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGTATTGGCTTCACTCA	322	
Db	241	TGTCTCAAGAGCTGGAATGGCCATTGGAATGGCGCTTACAGGAGTATTGGCTTCACTCA	300	
QY	323	CCACAACTGTACACTGTAATCCAGGCGCAATGAAGTTGGAAAGTCAAAATATTAGACA	382	
Db	301	CCACAACTGTACACTGTAATCCAGGCGCAATGAAGTTGGAAAGTAAAGAAATATTAGACA	360	
QY	383	GGGATTCATCACAGACCCCTGGTCTCTACAGCCCAAGAGATGGCTGGGATTTTGTGA	442	
Db	361	GGGATTCATCACTACTCTGTAGTCTCTTACCCCAAGATCGTGTAGGAGATTTTGTGA	420	
QY	443	GGCCAAAGCCCGGCGATGGTCTCTCGGTATCCCAATCACAGACAGGCGGATGGGGAG	502	
Db	421	AGCCAAAGCCAGGATGGCTTCTGTGTATCCCATCACAGATACAGGCGGATGGGGAG	480	
QY	503	CCGTTGGTGGCATCATCTCTCCAGAGCAATTGATTTTTCAAAGAGAGAGCAATCA	562	
Db	481	TGCGACTGGTGGCATCATCTCTCCAGAGCAATTGATTTTTCAAAGAGAGAGCAATCA	540	
QY	563	CTGTTCTTGGAGAGATTAATGACAAAGAGAGGAAGCTTGCTGTAGCCCGCCGCGCAT	622	
Db	541	CGGTTCTTGGAGAGATTAATGACAAAGAGAGGAAGTTGTGTGTGGCCCTGCAAGCAT	600	
QY	623	CACACTGAAGAGGCAATGAATTCGACGCGCAGCAAGAAAGGAAGTTGCCCATTTGT	682	
Db	601	CACCTGTGAAGAGGCAATGAATTTGACGCGCAGTAAAAAGGAAAGTTGCCCATTTGT	660	
QY	683	AAATGAATATGATGAGCTTGTGGGCATCATTTGCCCGGACAGACTGTAGAGAAATGGGA	742	
Db	661	GAATGAATATGATGAGCTGTGGGCATCATTTGCCCGGACAGACTGTAGAGAAATGGGA	720	
QY	743	CTACCCACTTACCCCTCCAAAGATGCGCAAGAAACAGCTGTGTGGGGCGAGCATTTGGAC	802	
Db	721	TTACCCATTTGGCTTCCAAAGATGCGCAAAAGCAGCTATATGTGGGGCAGCCATTGGTAC	780	
QY	803	TCATGAGGATGACAAGTATAGGCTTGACTTGCCTCGCCACAGGCTGGTGTGATGTAGTGT	862	
Db	781	TCATGAGGATGACAAGTATAGGCTGTGACTTACGCTCTTGTGGTGTGATGTAGTGTGT	840	
QY	863	TTTGGACCTTCCCGAGGAAATTCATCTTCCGATATCATATGATCAATATACATCAAGA	922	
Db	841	TTTGGACCTTCCCGAGGAAATTCATTTTCCAAATCATATGATCAATATACATCAAGA	900	
QY	923	CAAAATACCTTATCTCCAAAGTCATTTGGAGGCAATGTGTACTGCTGCCAGGCCAAGAA	982	
Db	901	GAATATACCTTATCTCCAAAGTCATTTGGAGGCAATGTATGACTGCTGCCAGGCCAAGAA	960	
QY	983	CTCATTTATGACAGTGTGGATGGCCCGGCGGAGGAGTGGGAAGTGGCTCATCTGTAT	1042	
Db	961	CTCATTTATGACAGTGTGTGATGTCTGCGAGTTGGCATATGGGGTGTGTTCTCATCTGTAT	1020	
QY	1043	TATCCAGCAAGCTGTGGCTGTGTGGCGGCGCCCAAGCAAGCAAGTGTACAGGTATATGA	11020	



	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
OY	743	CTACCCACTGAGCCTCCCAAGAATGCCAAGAAACAGCGTCGTGGGGGCAGCCATTGGCAC	M98333.1	GI:425157	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	1 (bases 1 to 1545)
Dd	721	TTACCATTGGCTTCCAAAGATGCCAAAAGCAGCTACTATGTGGGGCAGCCATTGGTAC					
OY	803	TCATGAGATGACAAGTATAGCGTGACTTGTCCGCCAGCGTGTTGCAATGATGTGT					
Dd	781	TCATGAGATGACAAGTATAGCGTGACTTGTCCGCCAGCGTGTTGCAATGATGTGT					
OY	863	TTTGAACCTCTTCCCAGGGAATTCATCTTCCAGATCAATATGATCAATACATCAAGA					
Dd	841	TTTTGACCTCTTCCCAGGGAATTCATCTTCCAGATCAATATGATCAATACATCAAGA					
OY	923	CAAAATACCCCTAATCTCCAAGTCATTTGAGAGCATGTGGCCTGCTCCAGCCAGAA					
Dd	901	GAAATATACCCCATCTTCCAAGTCATTTGAGAGCATGTGGCCTGCTCCAGCCAGAA					
OY	983	CCTCATTTGATGACAGTGTGGATGATCCCTCGGGGTGGCATGAGAGTGTTCATCTCAT					
Dd	961	CCTCATTTGATGACAGTGTGGATGATCCCTCGGGGTGGCATGAGAGTGTTCATCTCAT					
OY	1043	TATTCAGGAAGTGTGGGCTGTGGGGGGCCCCAACCAACAGAGTGTATAAGGTATGA					
Dd	1021	TACTAGGAAGTGTGGGCTGTGGGGGGCCCCAACCAACAGAGTGTATAAGGTATGA					
OY	1103	GTATGACAGCGCCCTTTGGTGTCCGGTCATTTCTCATGAGAGTAATCCAAATGTGGTCA					
Dd	1081	GTATGACAGCGCCCTTTGGTGTCCGGTCATTTCTCATGAGAGTAATCCAAATGTGGTCA					
OY	1163	TATTCAGGAAGCTTTGGCCCTTGGGGCCCTCCACAGTCATGATGGGTGTCTCTGGCTGC					
Dd	1141	TATTCAGGAAGCTTTGGCCCTTGGGGCCCTCCACAGTCATGATGGGTGTCTCTGGCTGC					
OY	1223	CACCACTGAGGCCCTGGTGAATACTTCTTCCGATGGATTCGGCTTAAGAAATATCG					
Dd	1201	CACCACTGAGGCCCTGGTGAATACTTCTTCCGATGGATTCGGCTTAAGAAATATCG					
OY	1283	CGGTATGGGTTCTTCGATGCCATGAGACAGCACCCTCAGACCCAGATAGATATTTCAG					
Dd	1261	CGGTATGGGTTCTTCGATGCCATGAGACAGCACCCTCAGACCCAGATAGATATTTCAG					
OY	1343	TGAACCTGACAAATCAAAAGTGCCCGCAGAGATGTCTGTGTGCGAGCAAAAGGGTC					
Dd	1321	TGAACCTGACAAATCAAAAGTGCCCGCAGAGATGTCTGTGTGCGAGCAAAAGGGTC					
OY	1403	AATCCACAATTTGTCCCTTACCTGATTTGCTGGATCCCAACACTGATCCAGAGATTGG					
Dd	1381	AATCCACAATTTGTCCCTTACCTGATTTGCTGGATCCCAACACTGATCCAGAGATTGG					
OY	1463	TGCCAAGAGCTTGACCCCAAGTCCGAGCCATGATGATCTGTGGGAGCTTAAGTTTGAGA					
Dd	1441	TGCCAAGAGCTTGACCCCAAGTCCGAGCCATGATGATCTGTGGGAGCTTAAGTTTGAGA					
OY	1523	GAGAACGTCTTCAGCCAGGTGGAGGTGGCTCCATAGAGCTTCATTTGTATGAGAAGG					
Dd	1501	GAGAACGTCTTCAGCCAGGTGGAGGTGGCTCCATAGAGCTTCATTTGTATGAGAAGG					
OY	1583	GCTTTTCTGAAAAGGATCCAGACACCTCTCTGGTTTTTTTTTCAA					
Dd	1561	GCTTTTCTGAAAAGGATCCAGACACCTCTCTGGTTTTTTTTTCAA					
RESULT 13							
MUSIMPD							
MUSIMPD							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							

FEATURES	Source	Location/Qualifiers
AUTHORS	Lightfoot, T. and Snyder, F.F.	
TITLE	Gene amplification and dual point mutations of mouse IMP	
JOURNAL	Biochim. Biophys. Acta 1217 (2), 156-162 (1994)	
MEDLINE	9415391	
PUBMED	7906545	
gene		1..1545 /organism="Mus musculus" /db_xref="taxon:10090" /cell_type="neuroblastoma" /tissue_type="brain" 1..1545 /gene="IMP" 1..1545 /gene="IMP" /codon_start=1 /product="IMP dehydrogenase" /protein_id="AAA20181.1" /db_xref="GI:425158" /translation="MADYLLISGTSYVPDDGLTAQQLFNCGDGLTYNDFLLPGYIDF TADVDLTSALTKITLKTPLVSSPMQTVTEAGMALMALTTGGIGFIHNCTPEFOAN EVKRVKKYEDGFIIDPVLVLSKRDVRDVFPAKRNHFGCIPIITDQRMGSRVGI1ISS RDIDFLKEEHDRLFEIIMTKREDLVVAPAGVTLFANELLORSKKGKPLVNEDEL VAIATRDLDKKNRDIPLASDKAKKOLLGSAIGTHEEDDKRLDLALAGDVVYVLDSS OGSNFIQIMTKYIKYKEXPSQVIGGVVTAQAQKNLIDGVDLRVMKSSGSICITD EYLAGCRPQATAVKVEYETARRGEPVYIADGGIDNVGH1AKALALGATVMSGLLAA TTEPAGYEPFSDGIRLRYKMGSLDMXKHLSSQNYFSEADIKVAQGVSVAVDQK GSIHKFPVYLIAGI0HSCODIGAKSLTQVBRAMYSGELKFEKRTSSAOVGVGSILHS YEKRLF" 998 /gene="IMP" /note="t in mutation" 1052 /gene="IMP" /note="a in mutation" 1052
variation		
variation		
BASE COUNT	390 a 374 c 419 g 362 t	
ORIGIN		
Query Match	78.5%: Score 1298.6; DB 10; Length 1545;	
Best Local Similarity	90.0%: Pred. No. 0;	
Matches 1391; Conservative	0; Mismatches 154; Indels 0; Gaps 0;	
QY	48 ATGGCCGACTACCTGATTAGTGGGGGACGTCCTACGTGCACAGAGAGGACTACACAGA	107
Db	1 ATGGCGGACTACCTGATTAGGAGGAGGACGCTTACGTGCGGAGAGCGGCTCACACGG	60
OY	108 CAGCAGCTCTTCACTGGCGAGAGCGGCTCACCTACCTACATGACTTTCATTTCTCCCTGGG	167
Db	61 CAGCAGCTCTCAACTGGGGGAGCGGCTCACAACATGATTTTCTCAATCTTCTCTGGG	120
OY	168 TACATCGACTCACGTGCGAGCAGGTGACCTGACCTTGTGCTGACCAACCAAAATTCCT	227
Db	121 TATATCGACTTCACGTGCGAGCAGGTGACCTGACCTTGTGCTTACTTAAAGATTACA	180
OY	228 CTTAAGACCCCACTGCTTCTCTCCCATGGACACAGTCACAGAGGCTGGATGGCCATA	287
Db	181 CTAAAGACCCCACTGCTTCTCTCCCACTGACACTCTTCACAGAGGCTGGATGGCCATC	240
OY	288 GCAATGGGCGCTTACAGCGGGGATTGGCTTCATCCACCACTGACACTGGAATTCAG	347
Db	241 GCGATGGGCGCTTACAGCGGGGATTGGCTTCATCCACCACTGACACTGGAATTCAG	300
OY	348 GCCAATGAAGTTGGAAAGTGAAGAAATATGAACAGGAGATTTCATCAGAGCCTGTGTC	407
Db	301 GCCAATGAAGTTGGAAAGTGAAGAAATATGAACAGGAGATTTCATCAGAGCCTGTGTC	360
OY	408 CTCAGCCCCAGAGATCGCGTCCGGATCTTTTGAAGCCAGAGCCCGGACATGTTTCGC	467
Db	361 CTTAGCCCCAGAGATCGGTACGGCATCTTTTGAAGCCAAAGCCAGGCAATGAGCTTCTGT	420
OY	468 GGATGCCCAATACAGACACAGGCGGATGGGAGAGCCCTTGTGTGGGATCATCTCTCC	527

Db	421	GGTATCCCCATCACAGTATACAGGCCGGATGGGAGTGCATTGGTGGCATCATCTCCCTCA	480
Qy	528	AGGGACATTGATTTTCTCAAGAGGAGGAACAATGACTGTGTTTCTTGGAGAGATATGACA	587
Db	481	AGGGACATTGATTTTCTCAAGAGGAGGAAGCATATGACCGGTCTTGGAGAGATCATGACT	540
Qy	588	AAGAGGACAGACTTGGTGTAGTACCCCCCGAGCATCACACTGAGAGGCAATGAATTT	647
Db	541	AAGAGGGAACATTTTGGTGTGCGCCCTCGCGGCGTCCACTCTCAAGAGAGCAATAGATTT	600
Qy	648	CTGAGGGCAGCAAGAGGGAAGTTGGCCATTGTAAATGAAGATGATGAGGTTGTGGCC	707
Db	601	CTGAGGGAGATAAAGGGAAGTTGGCCATTGTAAATGAAGATGATGAGTGGTATGCC	660
Qy	708	ATCATTTCCCGGAGACAGACTGGAAGAAATCGGGATACCCACTATGCCCTCCAAAGATGCC	767
Db	661	ATCATTTCCCGGAGACAGACTGGAAGAAATCGGTATGATACCCCTGGCTCCAAAGATGCC	720
Qy	768	AAGAAACAGCTGCTGTGTGGGGCAGCCATTGGCACTCATAGAGATGACAAATATAGCTG	827
Db	721	AAGAAACAGCTGCTGTGTGGGGCAGCCATTGGCACTCATAGAGATGACAAATATAGCTG	780
Qy	828	GACTTGTCCGCCAGGCTGGTGTGGATGTAGTGGTTTGGACTCTTCCAGGGAAATGCC	887
Db	781	GACTTGTCTGGCCCTTGTGTGTGGATGTAGTGGTTTGGACTCTTCCAGGGAAATGCC	840
Qy	888	ATCTTTCAGATCAATATATGATCAAGTACATCAAGCAAAATACCTATATCTCCAAATGCT	947
Db	841	ATCTTTCAGATCAATATATGATCAATATCAATCAAGAAATATCCCAAGTCTACAGSTCAT	900
Qy	948	GGAGGCAATGTGTCACTGCTGCTGCCAGGCCAGAACCTCATTTGATGACAGTGTGGATGCC	1007
Db	901	GGAGGCAATGTGTCACTGCTGCTGCCAGGCCAGAACCTCATTTGATGATGAGTGTAGATGCT	960
Qy	1008	CTGGGGGGGCGATGGGAAGTGGCTCCATCTGCATTATCCAGGAAGTCTGGCCCTGTGGG	1067
Db	961	TTGGGAGTGGCGGATGGGAAGTGGCTCCATCTGCATTATCCAGGAAGTGTGGCTGTGGG	1020
Qy	1068	CGGCCCCAAGCAACAGCAGTGTCAAGGTTATGATATGACAGCGCGCTTGGTGGTCCG	1127
Db	1021	CGGCCCCAAGCAACAGCAGTGTCAAGGTTATGATATGACAGCGCGCTTGGTGGTCTCT	1080
Qy	1128	GTCATTGCTGATGGAGGAATCAAAATGTGGGCTATATTGCGAAGAACCTTGGCCCTTGGG	1187
Db	1081	GTTATTGCTGATGGAGGAATCAAAATGTGGGCTATATTGCGAAGAACCTTGGCTTGGG	1140
Qy	1188	GGCTCCACAGCATGATGGGCTCTCTCTCGGGCTCCACACACTGAGGCCCCGTGTGAATAC	1247
Db	1141	GCTTCCACAGCATGATGGGCTCTCTCTCGGGCTCCACACACTGAGGCCCCGTGTGCAATAC	1200
Qy	1248	TTCTTTTTCGATGGAGTCCGGCTAAAGAAATATCGGGATGGGTTTCTCTCGATGCCATG	1307
Db	1201	TTCTTTTTCGATGGAGTCCGGCTGGAAGAAATATCCGAGGATATGGGTTTCTTGTATGCCATG	1260
Qy	1308	GACAAGCACTCAGCAGCCAGAAACAGATATTTTCATGTAAGCTGACAAATCAAACTGGCC	1367
Db	1261	GACAAGCACTCAGCAGCCAGAAACCATTACTTCACTGTAAGCTGACAAATCAAACTGGCC	1320
Qy	1368	CAGGAGATGTCTGCTGTGTGTCGAGGACAAAGGTCATATCCAAATTTGTCTCTTACTGT	1427
Db	1321	CAGGAGATGTCTGCTGTGTGTCGAGGACAAAGGTCATATCCAAATTTGTCTCTTACTGT	1380
Qy	1428	ATTGCTGGCATCCAAACCTCATGCCAGGACATTTGGTGGCAAGAGCTGACCCCAATGCCA	1487
Db	1381	ATTGCTGGCATCCAGCATTTCTCTCTGAGCATTTGGTGGCAAGAGCTTAACTCCAAATGCCA	1440
Qy	1488	GCCATGATGTACTTGTGGGAGCTTAAATTGAAAGACAGCACTCTCCAGCCAGGTGGAA	1547
Db	1441	GCCATGATGTACTTGTGGGAGGCTTAAATTTGAAAGACAGCACTCTCTGTCTCAGGTGGAA	1500
Qy	1548	GGTGGCGTCCATAGCTCCATTCTGTATGAGAAAGCGGCTTTTCTGA	1592

DB	1501	GGTGGCGTCCACAGCCTCCATTCGTACAGAGAAGCGCTTTCTGCA	1545
RESULT	14		
LOCUS	AK054667		
DEFINITION	Homo sapiens cDNA FLJ30105 fis. clone BNGH41000177, highly similar to INOS1N-5'-MONOPHOSPHATE DEHYDROGENASE 1 (EC 1.1.1.205).		PRI 01-AUG-2002
ACCESSION	AK054667		
VERSION	AK054667.1	GI:16549257	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens neuroglioma cell line:R4	CDNA to mRNA, clone_lib:BNGH41 clone:BNGH41000177.	
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Ito,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 2479)		
AUTHORS	Isogai,T., Otsuki,T. and Sugiyama,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory: 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.		
FEATURES			
Source	Location/Qualifiers		
	1..2479		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="BNGH41000177"		
	/cell_line="H4"		
	/cell_type="neuroglioma"		
	/clone_lib="BNGH41"		
	/note="Cloning vector: PME18SFL3"		
BASE COUNT	470 a 798 c 744 g 467 t		
ORIGIN			
Query Match	58.3%	Score 963.6	DB 9; Length 2479;
Best Local Similarity	76.5%	Pred. No. 51e-250;	
Matches 1182: Conservative	0; Mismatches 364; Indels	0; Gaps	0;
47	CATGGCGCACTCTGATGATGAGGGGCGACGCTCTACGCGCAGACGACGCTCACAGC	106	
DB	228 CATGGCGCACTCTGATGATGAGGGGCGACGCGGCTACGTCGCCGAGATGGCTTACGCC	287	
4Y	107 ACAGACAGCTCTTCAACTGCTGGAGAGGGGCTCACTACATGACTTTTCTTCATCTCCCTGG	166	
DB	288 GGAGCAGCTCTTGGCAGAGCGCGGAGGCGCTCACATCAACAGACTTCTGATCTCCGAG	347	
4Y	167 GTACATGCAGCTTACTGCAGACGACAGGTGAGCTGACTTCTGCTGTGACCAAGAAATCAG	226	
DB	348 ATTCAATAGACTTCATGATGATGATGATGATGAGCTGACCTGACCGCTGACCGGAAATCAG	407	
4Y	227 TCTTAAGACCCACGATGCTTCTCTCTCCATGAGACACAGTGCACAGAGGCTGGATGGCAT	286	
DB	408 GGTGAAGACGACCTGATCTCTCTCCCATGAGACACTGTGACAGAGGCTGATGATGGCAT	467	





```
/db_xref="GI:21706097"  
/db_xref="LocusID:3614"  
/translation="MADYLISGSGTYVEDGLTAQOLFASADGLYNDFLLIPGIDF  
IADEVDLTALTRKTLTKPLISSPMDVELEKMRGFGCIP1ETTGNGSKLVGTVS  
EVRKVKTEQGFITDPVYLSPSHTGMDVLEKMRGFGCIP1ETTGNGSKLVGTVS  
RDIDELAKRHHITLLEVTMPRIELVNPAGVTLEKAEIILQSRKGLVYNDEL  
VNIARTDLKKNRDYPLASDKSQKQLDGVAVGPNEDKRYDLTGQVGVTVLDS  
OGNSVYQIAMVHYIKOKYPHLOVIGNVTAQAANLIDAGVDLRYMGCGSICITQ  
EVMAGRGPGTAIVYVAEYARFVPIIADGLOTGVHVALAAGSTVMGSLIAA  
TTEARGEVFPDSQVRLKRRKRGMSIDAEKSSSOKRFESEGDRTQGVSGSTODK  
GSIQKFPYLLAGIHDGCDIGARSLSLVRKSMYSGELKFEKRTMSAIDEGVHLS  
TEKRLY"  
BASE COUNT      518 a      794 c      742 g      466 t  
ORIGIN  
Query Match      58.3% Score 963.6; DB 9; Length 2520;  
Best Local Similarity 76.5%; Pred. No. 5.1e-250;  
Matches 1182; Conservative 0; Mismatches 364; Indels 0; Gaps 0;  
47 CATGCCGCACTACCTGATTAGTGGGGGACGTCCTACGTGCAGAGACGAGCTCACAC 106  
||||| ||||||| ||| ||||| ||||||| ||| ||||| ||  
219 CATGGCGACTACCTGATTAGTGGGGGACCGGCTACGTGCCGAGAGATGGGCTACCGC 278  
||||| ||||||| ||| ||||| ||||||| ||| ||||| ||  
107 ACAGCAGCTCTTCACTGCGGAGAGCGCTCACCTAACATGACTTTCATTCCTCTGG 166  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
279 GCAGCAGCTCTTGCAGCGCGACGCGCTCACCTAACAGACTTCTGATTCCTCCAGG 338  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
167 GTACATCGACTTACGTCGACCGACCGTGCCTGCTGCTGACCAAGAAATAC 226  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
339 ATTCAATAGCTTCAATAGCTGATGAGTGAAGCTGACCTTACGCTTACCGGGAAGTAC 398  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
227 TCTTAAGACCCCACTGGTTTCTCCATGACACAGTACAGAGCGTGGATGGCAAT 286  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
399 GGTGAAGAGCGCACTGATCTCTCCCCCATGGACACTGACAGAGAGCGACATGGCCAT 458  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
287 AGCAATGGCGTTACAGCGGCTTATGCTTATCCACCACTGTTACCTGAAATTCGA 346  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
459 TGCATGAGCTTGTGATGGAGGATATGTTTTCATTCACCACTGACCCCGAGATTCGA 518  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
347 GGCATGTAAGTTCGGAAGTGAAGAAATATGACAGGGGATTCATCAGACGCTGTGGT 406  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
519 GGCACAGAGGTGCGAAGAGTCAAGAAATTTGACAGGGCTTCATACGACCGCTGTGGT 578  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
407 CCTCAGCCCCAAGATGCGGTGGCGGATGTTTGAAGCCAAAGCCCGCATGGTTCTG 466  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
579 GCTAGCCCCCTCGACACAGTGGGCGATGCTGGAGCGCAAGATCGGCGATGGCTCTC 638  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
467 CGGTATCCCAATACAGACAGCGCGGATGGGAGCCGCTTGGTGGCATCTCTCTC 526  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
639 TGGCATCCCATCTACAGAACGGGACGATGGGACGACGCTGGGATCGTCACTC 698  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
527 CAGGCACTTGAATTTCTAAAGAGGAGAACATGACTGTTTCTGGAAGATTAATGAC 586  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
699 CCGAGACATCGACTTCTTCTGTGAGAGAACACACACCCTCTCTAGTGGGTGTGAC 758  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
587 AAAGAGGAAGACTTGGTGTAGCCCCCGCAGACATCACACTGAAAGAGGCAAAATGAAAT 646  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
759 GCCAAGGATTGAACGTGGTGGCTCCACAGGTGTGACGTTGAAAGAGGCAAAATGAGAT 818  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
647 TCTGCAGCGCAGAGAAAGGAAAGTTGCCATTTGAATGAAGATGATGAGCTTGGC 706  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
819 CCTGCAGCGCTAGCAAGAAAGGAGAGCTGCTATGCTCAATGATTGCGATGAGCTGTGGC 878  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
707 CATCATTTGCCGAGACGACTGAAAGAAATGGGACTACCCACTAGCCCTCCAAAGATGC 766  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
879 CATCATGGCCCGACCGCTTCTGTAAGAGAAACGAGACTACCCCTGTGGCTCCAAAGGATTC 938  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
767 CAGAAGACAGCTGCTGTGGGCGACCATTTGGCACTCATGAGATGACAAAGTATAGCT 826  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
939 CAGAAGACAGCTGCTGTGGGCGACCTGTGGGCGACCCGTGAGAGATGACAAATACGCTCT 998  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||  
827 GGACTGTGCGCCAGAGCTGTGTGATGATGTTTGGACTCTTCCAGGGAATTC 886  
||||| ||||||| ||| ||||||| ||||||| ||||| ||||| ||
```

```
Db 999 GGACCTGCTCACCCAGCGGGCGGTGAGCTCATAGCTTTGACCTGCTCCAGGAATTC 1058  
887 CATCTTCAGATCATATGATCATAGTACATCAAGCAAAATACCTTAATCTCAGATCAT 946  
||| ||||| ||| ||||| ||||| ||| ||||| ||||| ||| ||||| ||  
1059 GGTATATCATGATCGCATGCTCATATTACATCAAGAGAAATACCCACCTCCAGTGAT 1118  
947 TGAAGCAATGTGTGCTACTGCTGCCAGGCCAAGAACCTCATTTGAGTGCAGGTGGATGC 1006  
||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1119 TGGGGGGAACGTGTGACAGCAGCCAGCCAGCAAGAACCTGATTTGATGCTGGTGGACGG 1178  
1007 CCTGGGGGTGGGCAATGGGAAGTGGCTCATCTGCAATTAATCCAGGAAGTGTGGCTTGG 1066  
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1179 GCTGGCGGTGGCATGGCTGGCTCCATTCGATCACACCAGGAAGTATGGCTGTGG 1238  
1067 GCGCCCCCAAGCAAGCAGCATGTACAGGTTGATAGTATGACGCGCTTGTGTGCTCC 1126  
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1239 TGGGCCCGGAGCACTGCTGTATACAGGTGCTGATGATGCGCGCTTTGGTGTGCG 1298  
1127 GGTCAATTCGTGATGAGCAATCAAAATGTGGTCAATATTCGAAAGACCTTGGCCCTTGG 1186  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1299 CATCATAGCCGATGGGGGATCCAGACCGTGGGACACGTGTCAAGGCCCTTGGCTTGG 1358  
1187 GGCCTTCACAGTATGATGGGCTCTCTCTGCTGCCACACTGAGGCGCCCTGTGATA 1246  
||||| ||||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1359 ACCCTTCACAGTATGATGGGCTCTCTGCTGCCACACTGAGGCGCCCTGTGAGATA 1418  
1247 CTCTCTTTCGATGGGATCCGGCTAAAGAAATATCGCGGTATGGGTCTCTGATGCCAT 1306  
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1419 CTCTCTTCTAGACGGGCTGCGCTCAAGAAATACCGGGGATATGGCTCACTGATGCCAT 1478  
1307 GACACACACCTCACAGCCAGCAAGATATTTCAAGTCAAGCTGACAAATCAAGATGGC 1366  
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1479 GGAAGAGACACAGACAGCAGCAAGAAAGATCTTCAAGCGGGGGGATTAAGTGAATGCG 1538  
1367 CCAGGAGTGTGTGCTGCTGTGACAGCAAGGATCAATCCCAATTTGTCTTACT 1426  
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1539 GAGAGTGTCTGGGCTCCATCCAGAACAAAGATTCATTCAGAAAGTTCGTCCCTACT 1598  
1427 GATTGCTGCATCCAAACACTATGCCAGGACATTGGTGCAGAGACTTGACCAAGTCCG 1486  
||| ||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1599 CATAGCAGCATCCAAACAGCGCTGCCAGATATCGGGGCCCGCAGCTGTCTCTTGC 1658  
1487 ACCATGATTACTCTGGGAGCTTAAGTTGAGAAAGAGAACCTCTCAGCCAGGTGGA 1546  
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1659 GTCCATGATGATCTCAGGAGACTCAAGTTGAGAAAGCGACCATGTGGCCAGATTGA 1718  
1547 AGGTGGCGTCCATGAGCTTCATTCGATGAGAAAGCGGCTTTCTGA 1592  
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
1719 GGTGTGTTCATGGCTGCTTACGAAAAAGCGGCTGTACTGA 1764
```

Search completed: February 12, 2003, 18:52:29  
Job time : 4429 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 17:32:54 (Search time 81 seconds  
(without alignments))

6652.264 Million cell updates/sec

Title: US-09-846-637C-3

Perfect score: 1654

Sequence: 1 gaatcgcgcgcgtcctcgcga.....gttagaagaccgcattc 1654

Scoring table: IDENTITY-NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*

2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*

3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*

4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*

5: /cgn2\_6/prodata/2/ina/PCRTUS\_COMB.seq:\*

6: /cgn2\_6/prodata/2/ina/backfile1.seq:

7: /cgn2\_6/prodata/2/ina/backfile1.seq:

8: /cgn2\_6/prodata/2/ina/backfile1.seq:

9: /cgn2\_6/prodata/2/ina/backfile1.seq:

10: /cgn2\_6/prodata/2/ina/backfile1.seq:

11: /cgn2\_6/prodata/2/ina/backfile1.seq:

12: /cgn2\_6/prodata/2/ina/backfile1.seq:

13: /cgn2\_6/prodata/2/ina/backfile1.seq:

14: /cgn2\_6/prodata/2/ina/backfile1.seq:

15: /cgn2\_6/prodata/2/ina/backfile1.seq:

16: /cgn2\_6/prodata/2/ina/backfile1.seq:

17: /cgn2\_6/prodata/2/ina/backfile1.seq:

18: /cgn2\_6/prodata/2/ina/backfile1.seq:

19: /cgn2\_6/prodata/2/ina/backfile1.seq:

20: /cgn2\_6/prodata/2/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1635.6	98.9	1642	3	US-08-925-230-1
2	1321.8	79.9	1620	3	US-08-925-230-2
3	420.6	25.4	675	4	US-09-328-111-844
4	410.2	24.8	3616	4	US-09-212-247C-7
5	285.8	17.3	393	3	US-08-925-230-5
6	285.8	16.7	344	3	US-08-925-230-6
7	275.8	16.7	344	3	US-08-925-230-3
8	275.8	16.7	344	3	US-08-925-230-4
9	226.2	13.7	3531	4	US-08-629-600-1
10	226.2	13.7	3531	4	US-09-076-761-1
11	179.4	10.8	1485	4	US-09-134-001C-2187
12	154.6	9.3	5884	4	US-09-221-017B-497
13	118.4	7.2	1653	4	US-08-961-527-329
14	81.6	4.9	1843	1	US-08-774-169-2
15	75.8	4.6	13121	4	US-08-961-527-126
16	71.8	4.3	1002	4	US-09-134-001C-2367
17	42.6	2.6	7218	1	US-08-232-463-14
18	40	2.4	309	4	US-09-134-001C-2158
19	38.8	2.3	5661	4	US-08-938-105-2
20	35.8	2.2	2353	1	US-08-328-322-11
21	35	2.1	7326	1	US-08-194-468-1
22	35	2.1	7326	4	US-09-514-247A-7
23	35	2.1	7344	3	US-08-961-739-1
24	35	2.1	8752	4	US-08-976-259-3
25	34.8	2.1	421	4	US-08-642-274D-55
26	34.8	2.1	421	4	US-08-952-014C-55
27	34.4	2.1	3072	4	US-09-522-217-55

C	28	33.2	2.0	5000	3	US-09-104-070-3	Sequence 3, Appl
C	29	32.6	2.0	1396	4	US-09-328-174A-2	Sequence 2, Appl
C	30	32.4	2.0	723	4	US-08-998-416-581	Sequence 581, App
C	31	32.4	2.0	2380	1	US-08-572-951-3	Sequence 3, Appl
C	32	32.2	1.9	1957	4	US-09-352-990-11	Sequence 11, Appl
C	33	32.2	1.9	4249	4	US-09-071-035-403	Sequence 403, App
C	34	32.2	1.9	4359	4	US-09-071-035-401	Sequence 401, App
C	35	32	1.9	954	3	US-08-911-853-24	Sequence 24, Appl
C	36	32	1.9	954	4	US-09-479-409-24	Sequence 24, Appl
C	37	32	1.9	954	4	US-09-479-453-24	Sequence 24, Appl
C	38	32	1.9	17612	3	US-08-911-853-29	Sequence 29, Appl
C	39	32	1.9	17612	4	US-09-479-409-29	Sequence 29, Appl
C	40	32	1.9	17612	4	US-09-479-453-29	Sequence 29, Appl
C	41	31.6	1.9	2057	3	US-09-008-303-1	Sequence 1, Appl
C	42	31.4	1.9	1332	4	US-09-134-001C-930	Sequence 930, App
C	43	31.4	1.9	3766	4	US-08-961-527-225	Sequence 225, App
C	44	31.4	1.9	4359	4	US-09-484-970B-4	Sequence 4, Appl
C	45	31.2	1.9	1296	1	US-08-391-339-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-925-230-1  
Sequence 1, Application US/08925230

Patent No. 6147194

GENERAL INFORMATION:

APPLICANT: Huberman, Eliezer

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS: Arnold, White & Durkee

ADDRESS: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/925,230

FILING DATE: September 8, 1997

CLASSIFICATION: 536

PRIOR APPLICATION NUMBER: US 5,665,583

ATTORNEY/AGENT INFORMATION:

NAME: Wilson, Mark B.

REGISTRATION NUMBER: 37,259

REFERENCE/DOCKET NUMBER: ARCD:274

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1642 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-925-230-1

Query Match

Best Local Similarity 98.9%

Matches 1658: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

7 GGGCGGCTCTCGACACCGCGGCTGCTGTGGCCATGCGCCGACTGATTA 66

Db 1 GGGCGGTCCTCGAGACACAGCGGGGTGCTGTGTGGCCATGGCCGACTACTGTGATTA 60  
Qy 67 GTGGGGGCGACGTCCTACCGTGGCAGACGAGCTCAGCAGCAGCAGCTTCAACTGCG 126  
Db 61 GTGGGGGCGACGTCCTACCGTGGCAGACGAGCTCAGCAGCAGCAGCTTCAACTGCG 120  
Qy 127 GAGAGGGGCTTCACATCAATGACTTTCATTCCTCCCTGGGTACATGACTTCACTGCG 186  
Db 121 GAGAGGGGCTTCACATCAATGACTTTCATTCCTCCCTGGGTACATGACTTCACTGCG 180  
Qy 187 ACCAGGTGACCTGACTTTCGCTGTGACCAAGAAATTCATCTTAAAGACCCACGCTTT 246  
Db 181 ACCAGGTGACCTGACTTTCGCTGTGACCAAGAAATTCATCTTAAAGACCCACGCTTT 240  
Qy 247 CCTCTCCCATGGACACAGTGCAGAGGCTGGATGGCCTATAGCAATGGCGCTTACAGCG 306  
Db 241 CCTCTCCCATGGACACAGTGCAGAGGCTGGATGGCCTATAGCAATGGCGCTTACAGCG 300  
Qy 307 GTATTGGCTTTCATCCACCAACTGTACACTGAAATTCAGGGCCAATGAAATTCGGAAG 366  
Db 301 GTATTGGCTTTCATCCACCAACTGTACACTGAAATTCAGGGCCAATGAAATTCGGAAG 360  
Qy 367 TGAAGAAATATGACAGGGGATTCATCAGACCCCTGGGCTCAGCCCAAGGATCGCG 426  
Db 361 TGAAGAAATATGACAGGGGATTCATCAGACCCCTGGGCTCAGCCCAAGGATCGCG 420  
Qy 427 TCGGGATGTTTTTGAAGCCAAAGCCCGCATGGTTCTGCGGTATCCCAATCACAGACA 486  
Db 421 TCGGGATGTTTTTGAAGCCAAAGCCCGCATGGTTCTGCGGTATCCCAATCACAGACA 480  
Qy 487 CAGCGCGGATGGGGAGCCGCTTGTGGGCATCTCTCTCCAGGGCATTGATTTTCTCA 546  
Db 481 CAGCGCGGATGGGGAGCCGCTTGTGGGCATCTCTCTCCAGGGCATTGATTTTCTCA 540  
Qy 547 AAGAGAGAGAACATGACTGTTCTTGAAGAGATTAATGACAAAGAGGAGGACTTGTG 606  
Db 541 AAGAGAGAGAACATGACTGTTCTTGAAGAGATTAATGACAAAGAGGAGGACTTGTG 600  
Qy 607 TAGCCCCCGCAGCATCAGACTGAAAGAGGCAATGAATTTCTGACGCGACGCAAGAG 666  
Db 601 TAGCCCCCGCAGCATCAGACTGAAAGAGGCAATGAATTTCTGACGCGACGCAAGAG 660  
Qy 667 GAAAGTTGCCATTTGAATGAAGATGATGAGCTTGTGGCATATTTGCCCGACAGACC 726  
Db 661 GAAAGTTGCCATTTGAATGAAGATGATGAGCTTGTGGCATATTTGCCCGACAGACC 720  
Qy 727 TGAAGAAAGATCGGAGTACCCACTAGCCCTCCAAAGATGCGCAAGAAACAGCTGTG 786  
Db 721 TGAAGAAAGATCGGAGTACCCACTAGCCCTCCAAAGATGCGCAAGAAACAGCTGTG 780  
Qy 787 GGGCAGCCCATTTGGCACTCATGAGATGACAAGTATAGGCTGACTTGTCCCGCAGGCT 846  
Db 781 GGGCAGCCCATTTGGCACTCATGAGATGACAAGTATAGGCTGACTTGTCCCGCAGGCT 840  
Qy 847 GTGTGATGTAGTGTGTTTTGGACTCTTCCAGGGAAATTCATCTTCCAGATCAATATGA 906  
Db 841 GTGTGATGTAGTGTGTTTTGGACTCTTCCAGGGAAATTCATCTTCCAGATCAATATGA 900  
Qy 907 TCAAGTACATCAAAAGCAAAATACCTAATCTCCAAAGTATTTGAGGCAATGTGCTACT 966  
Db 901 TCAAGTACATCAAAAGCAAAATACCTAATCTCCAAAGTATTTGAGGCAATGTGCTACT 960  
Qy 967 CTGCCCCAGGCAAGAACTCATTTGATGACAGTGTGATGCTGCGGGTGCGGATGGAA 1026  
Db 961 CTGCCCCAGGCAAGAACTCATTTGATGACAGTGTGATGCTGCGGGTGCGGATGGAA 1020  
Qy 1027 GTGGCTTCATCTGATTAATTCAGGAAGTGTGCTGTGGCGGGCCCAAGCAACAGCAG 1086  
Db 1021 GTGGCTTCATCTGATTAATTCAGGAAGTGTGCTGTGGCGGGCCCAAGCAACAGCAG 1080  
Qy 1087 TGTACAAGGTGTGATGATGACAGGCGCTTGTGTGCTCCGGTCAATTTGCGATGGAGAA 1146  
Db 1081 TGTACAAGGTGTGATGATGACAGGCGCTTGTGTGCTCCGGTCAATTTGCGATGGAGAA 1140

Qy 1147 TCCAAATGTGGGTATATTTGCGAAAGCCCTTGGCCCTTGGGCTCCACAGTCAATGATG 1206  
Db 1141 TCCAAATGTGGGTATATTTGCGAAAGCCCTTGGCCCTTGGGCTCCACAGTCAATGATG 1200  
Qy 1207 GCTCTCTCTGGCTGCGACACTAGAGCCCCCTGGTGAATACTTCTTTCCGATGGGATCC 1266  
Db 1201 GCTCTCTCTGGCTGCGACACTAGAGCCCCCTGGTGAATACTTCTTTCCGATGGGATCC 1260  
Qy 1267 GCGTAAAGAAATATCGCGGTATGGTTCTCTGATGCGCATGGACAGCACTCAGACGC 1326  
Db 1261 GCGTAAAGAAATATCGCGGTATGGTTCTCTGATGCGCATGGACAGCACTCAGACGC 1320  
Qy 1327 AGACAGATATTTTCAGTGAAGCTGACAAATCAAAAGTGGCCCAAGGAGTGTGCTG 1386  
Db 1321 AGACAGATATTTTCAGTGAAGCTGACAAATCAAAAGTGGCCCAAGGAGTGTGCTG 1380  
Qy 1387 TGCAGACAAAGGTCATTCACAAATTTGCTCCTTACCTGATTTGCTGGCATCCAACT 1446  
Db 1381 TGCAGACAAAGGTCATTCACAAATTTGCTCCTTACCTGATTTGCTGGCATCCAACT 1440  
Qy 1447 CATGCCAGACATTTGTGTCGAAGAGCTTGACCCCAAGTCCGAGCCATGATGTACTG 1506  
Db 1441 CATGCCAGACATTTGTGTCGAAGAGCTTGACCCCAAGTCCGAGCCATGATGTACTG 1500  
Qy 1507 AGCTTAAGTTTGAAGAGACAGTCTCTCAGCCAGGTGGAAGGTGGCTTCATAGCTCC 1566  
Db 1501 AGCTTAAGTTTGAAGAGACAGTCTCTCAGCCAGGTGGAAGGTGGCTTCATAGCTCC 1560  
Qy 1567 ATTGCTATGAGAAGGCGCTTCTGTGAAGAGGATCCAGCACACCTCTCGGTTTTTTT 1626  
Db 1561 ATTGCTATGAGAAGGCGCTTCTGTGAAGAGGATCCAGCACACCTCTCGGTTTTTTT 1620  
Qy 1627 CAATAAAGTTTGAAGAAGCC 1648  
Db 1621 CAATAAAGTTTGAAGAAGCC 1642

RESULT 2  
US-09-925-230-2  
; Sequence 2, Application US/08925230  
; Patent No. 6147194  
; GENERAL INFORMATION:  
; APPLICANT: Coliart, Frank  
; APPLICANT: Huberman, Eliezer  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH  
; TITLE OF INVENTION: AND GMP PRODUCTION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/925,230  
; FILING DATE: September 8, 1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 5,665,583  
; FILING DATE: 12-AUG-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mark B.  
; REGISTRATION NUMBER: 37,259  
; REFERENCE/DOCKET NUMBER: ARCD:274  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1620 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-925-230-2

Query Match Best Local Similarity 79.9% Score 1321.8; DB 3; Length 1620;  
 Matches 1442; Conservative 0; Mismatches 177; Indels 4; Gaps 1;

23 CACGCGGCGGTCTCTGTGTTGGCCATGGCCGCACTACTGATTTAGTGGGCGACGCTCTA 82  
 1 CACGCGTCCGCTCTCTGTCGTCATGGCCATGGCGGACTACTGATTTAGGCGGCGACATCTTA 60  
 83 CGTGGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 142  
 61 CGTGGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120  
 143 CAATGACTTCTCATTTCTCTGCTGGGACATGACGACGACGACGACGACGACGACGACGACGAC 202  
 121 CACGATTTCTCATTTCTCTGCTGGGACATGACGACGACGACGACGACGACGACGACGACGAC 180  
 203 TTCTGCTGACCAAGAAATCACTCTTAAGACCCCACTGGTTTCTCT 160  
 181 CTCTGCTTCACTCAATCAATCACTCTTAAGACCCCACTGGTTTCTCT 140  
 263 ACTGACAGAGGCTGGGATGGCCATGACGACGACGACGACGACGACGACGACGACGACGACGAC 322  
 241 TGTGACAGAGGCTGGGATGGCCATGACGACGACGACGACGACGACGACGACGACGACGACGAC 300  
 323 CCACAACTGTAACCTGATTTCCAGGCAATGAGTTGGGAAGTGA 382  
 301 CCACAACTGTAACCTGATTTCCAGGCAATGAGTTGGGAAGTGA 360  
 383 GCGATTCATCAGAGACCCCTGCTCTGACCCCAAGAGATCCGCTGCGGATGTTTGA 442  
 361 GCGATTCATCAGAGACCCCTGCTCTGACCCCAAGAGATCCGCTGCGGATGTTTGA 420  
 443 GCGCAAGCCCGGATGTTTCTGCGGATTTCCCAATCAGACACAGCCCGGATGGGAG 502  
 421 ACCCAAGCCCGGATGTTTCTGCGGATTTCCCAATCAGACACAGCCCGGATGGGAG 480  
 503 CCGCTTGGTGGGATGATCTCTCCAGGACATGATTTTCCAAAGAGAGAGACATGA 562  
 481 TCGACGCGTGGGATGATCTCTCCAGGACATGATTTTCCAAAGAGAGAGACATGA 540  
 563 CTGTTTCTTGAAGATATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 622  
 541 CCGTTTCTTGAAGATATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 600  
 623 CACACTGAAGAT 682  
 601 CACTCTGAAGAT 660  
 683 AAATGAAGATGATGACCTTGTGGCCATCATTTGCCGACAGACCTGTA 742  
 661 GAATGAAGATGATGACCTTGTGGCCATCATTTGCCGACAGACCTGTA 720  
 743 CTACCCACTAGGCTCCAAAGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 802  
 721 TTACCCATTTGGCTTCCAAAGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780  
 803 TCATGAGATGACCAAGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 862  
 781 TCATGAGATGACCAAGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840  
 863 TTGTCGCTCTCCAGAT 922  
 841 TTGTCGCTCTCCAGAT 900

923 CAAATACCCCTAATCTCAAGTCAATTTGAGGCAATGAGTGGTCACTGGTGGCCAGGCAAGAA 982  
 901 GAAATACCCCAATCTCAAGTCAATTTGAGGCAATGAGTGGTGGTCACTGGTGGCCAGGCAAGAA 960  
 983 CCTCATTTGATGACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042  
 961 CCTCATTTGATGACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 1043 TATCCAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1102  
 1021 TACTCAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080  
 1103 GTATGACGCGCCCTTTGTTGTTCCGCTCATTTGCTGATGAGAGAAATCCAAATGAGTGA 1162  
 1081 GTATGCTCGCGCTTTGTTGTTCCGCTCATTTGCTGATGAGAGAAATCCAAATGAGTGA 1140  
 1163 TATGCGAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1222  
 1141 TATGCGAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1200  
 1223 CACCACTGAGGCGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1282  
 1201 CACCACTGAGGCGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1260  
 1283 CCGTATGGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342  
 1261 TGGTATGGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 1343 TGAAGCTGCAAAATCAAAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1402  
 1321 TGAAGCTGCAAAATCAAAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
 1403 AATCCAAATTTGTCCTTACCCTGATGCTGAGATCCAGACACTCATGCGCAGAGACTGG 1462  
 1381 TATCCAAATTTGTCCTTACCCTGATGCTGAGATCCAGACACTCATGCGCAGAGACTGG 1440  
 1463 TGCCAAGAGCTTGACCCCAAGTCCGACCATGATGATGATGATGATGATGATGATGATGATGAT 1522  
 1441 TGCCAAGAGCTTGACCCCAAGTCCGACCATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1523 GAGAGCTGCTCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1582  
 1501 GAGAGCTGCTCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
 1583 GCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCAATTAAGATTAGAA 1642  
 1561 GCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCAATTAAGATTAGAA 1620  
 1643 AGA 1645  
 1617 AAA 1619

RESULT 3  
 US-09-328-111-844  
 ; Sequence 844, Application US/09328111  
 ; Patent No. 6262333  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Endege, Wilson O.  
 ; APPLICANT: Steinmann, Kathleen E.  
 ; APPLICANT: Astle, Jon H.  
 ; APPLICANT: Burgess, Christopher C.  
 ; APPLICANT: Bushnell, Steven E.  
 ; APPLICANT: Carroll III, Eddie  
 ; APPLICANT: Catlino, Theodore J.  
 ; APPLICANT: Derfl, Adnan  
 ; APPLICANT: Ford, Donna M.  
 ; APPLICANT: Lewis, Marcia E.  
 ; APPLICANT: Monahan, John E.  
 ; APPLICANT: Schlegel, Robert  
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
 ; TITLE OF INVENTION: PRODUCTS  
 ; FILE REFERENCE: CCD-257 (US)

;; CURRENT APPLICATION NUMBER: US/09/328,111  
;; CURRENT FILING DATE: 1999-06-08  
;; EARLIER APPLICATION NUMBER: US 60/088,801  
;; EARLIER FILING DATE: 1998-06-10  
;; NUMBER OF SEQ ID NOS: 850  
;; SOFTWARE: FASTSEQ for Windows Version 3.0  
;; SEQ ID NO 844  
;; LENGTH: 675  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(675)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-328-111-844

Query Match 25.4%; Score 420.6; DB 4; Length 675;  
Best Local Similarity 95.8%; Pred. No. 2,3e-120;  
Matches 432; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 331 GTACACCTGAATTCAGGCCAATGAAGTTCGGAAGTAATATGAACAGGATTC 390  
DB 2 GTACACCTGAATTCAGGCCAATGAAGTTCGGAAGTAATATGAACAGGATTC 61  
OY 391 TCACAGACCCCTGTGCTCTCAGCCCAAGATCGCTGGGGATGTTTTTGAAGCCCAAG 450  
DB 62 TCACAGACCCCTGTGCTCTCAGCCCAAGATCGCTGGGGATGTTTTTGAAGCCCAAG 121  
OY 451 CCCGATGATGTTTCTGCGGTATCCCAATCAGACACAGCCCGGATGGGAGCCGTTGG 510  
DB 122 CCCGATGATGTTTCTGCGGTATCCCAATCAGACACAGCCCGGATGGGAGCCGTTGG 181  
OY 511 TGGGATCATCTCTCCAGGACATGATTTTCTCAAGAGAGGAACATGACTGTTCT 570  
DB 182 TGGGATCATCTCTCCAGGACATGATTTTCTCAAGAGAGGAACATGACTGTTCT 241  
OY 571 TGGAGAGATATGACAAAGAGGAAAGCTTGGTGTAGCCCCCGACATCACACTGA 630  
DB 242 TGGAGAGATATGACAAAGAGGAAAGCTTGGTGTAGCCCCCGACATCACACTGA 301  
OY 631 AGGAGCAATGAATTCGACGCGCAGCAAGAAGGAAGTGGCCATTTGTAATGAAG 690  
DB 302 AGGAGCAATGAATTCGACGCGCAGCAAGAAGGAAGTGGCCATTTGTAATGAAG 361  
OY 691 ATGATGAGCTTGGCCATCATTTGCCGAGACACTGAAGAAGATTCGGATACCCAC 750  
DB 362 ATGATGAGCTTGGCCATCATTTGCCGAGACACTGAAGAAGATTCGGATACCCAC 421  
OY 751 TAGCTCCAAAGATGCCAAGAAACAGCTGCT 781  
DB 422 TAGCTCCAAAGATGCCAAGAAACAGCT 452

RESULT 4  
US-09-212-247C-7  
; Sequence 7, Application US/09212247C  
; Patent No. 6391603  
; GENERAL INFORMATION:  
; APPLICANT: POMPEJUS, Markus; SUELBERGER, Harald; JOEPFEN, Hans  
; WOLFGANG, DOVAL, Jose Luis Revuelta; JIMENEZ, Alberto;  
; and GARCIA, Maria Angeles Santos  
; TITLE OF INVENTION: Genes of purine biosynthesis from Asbhya Gossypii  
; and the use thereof in microbial riboflavin  
; synthesis  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: 1101 Connecticut Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
;; COMPUTER: IBM AT-compatible, Pentium processor  
;; OPERATING SYSTEM: Windows 95  
;; SOFTWARE: Wordperfect version 6.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/212,247C  
;; FILING DATE: 16-Dec-1998  
;; CLASSIFICATION: <Unknown>  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3616 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..863  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 864..1316  
;; FEATURE:  
;; NAME/KEY: Intron  
;; LOCATION: 1317..1477  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1478..2592  
;; FEATURE:  
;; NAME/KEY: 3'UTR  
;; LOCATION: 2593..3616  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-212-247C-7

Query Match 24.8%; Score 410.2; DB 4; Length 3616;  
Best Local Similarity 62.2%; Pred. No. 1.1e-116;  
Matches 709; Conservative 0; Mismatches 403; Indels 27; Gaps 3;

OY 465 TGGGATATCCCAATGACACACAGCCCGGATGGGAGCCGCTGGGATCATCTGCC 524  
DB 1460 TGTACTACATCATGACACATGATGGCAAGCCGAGCGGAGACTGCAGGGATCATCAG 1519  
OY 525 TCCAGGACATTTGTTCTCAAGAGAGAAACATGACTGTTCTTGAAGAGATATG 584  
DB 1520 TCCCGATCATCAGTTGTCGAGCAGCAG-----ACCTGCTGTGCTGTAG 1567  
OY 585 ACAAGAGGGAAGACTTGTGTGTAGCCCCCGCAGCATCACACTGAAGAGCAATGAA 644  
DB 1568 ATCATGACCCAAAGAGCTCATCTGTGGAAGCAGGGCATCACTGAGAGGCGAACCAG 1627  
OY 645 ATTCTGACGCGAGCAGAGAGGAAGTGGCCATTTGTAATGAAGATGAGCTGTG 704  
DB 1628 ATCTGAGAACACCAAGAGGAGCAGCTGCCAATTTGTGACAGAGCGCGCTGCTGTG 1687  
OY 705 GCCATCATTTGCCCGACAGACTGAAGAAGATCGGACTACCACTAGCTCCAAAGAT 764  
DB 1688 TCCATGCTTTGAGAACTGATGATGAAGAACAGTCTTACCATTGGCTTCCAAAGTCT 1747  
OY 765 GCCAA---GAACAGCTGCTGTGTGGGACACCATTTGCACATGAGATGACAAATAT 821  
DB 1748 GCCGACACCAACAGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1807  
OY 822 AGGCTGACTTGTGCTGCCAGGCTGTGTGATGATGATGATGATGATGATGATGATGAT 881  
DB 1808 AGACTGGCATGCTGTGAGGCGGCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1867  
OY 882 AATTCCATCTTCAGATCATATGATCAATCAATCAAGAAATACCTAATCTCCAA 941  
DB 1868 AACTGGTCTTCCAGATCAATGATCAATGATCAATGATCAATGATCAATGATCAATGAT 1927  
OY 942 GTCAATTTGAGGCAATGTGTGTCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001

```

Db 1928 GTCATGCTGGCAGCGTGTCAACAGAGCAAGCGTCCAGCTTGATCAACGCCGGCGCA 1987
Qy 1002 GATGCCCTCGGGGTGGGATGGGAAGTGGCTCCATTCGATTAATCCAGAAAGTGGTGGCC 1061
Db 1988 GACGGGTTGCGATGATGGCTCTGGCTCCATCTGATCACTAGAGGTGATGGCC 2047
Qy 1062 TGTGGCGGCCCCAAGCAGCAGCAGTGTACAAAGGTGTATGAGATGCAAGCGCTTGGT 1121
Db 2048 TGTGTAAGACACAGAGGTACCGCTGTCTACAAAGTACAGCAGTGGCCAAACGTTTGGT 2107
Qy 1122 GTTCCGGTCTTCTGATGAGAGAAATCCAAATGTGGTCAATATTCGCAAGCGTTGGCC 1181
Db 2108 GTGCCATGTATTGTGACCGGTGTCCAGAACATCGGCACATTCACAAGCTATCGCT 2167
Qy 1182 GTTGGGCGCTCCAGCAGTCAATGATGGCTCTCTGCTGGCGCCAGCAGCTAAGCGCTGGT 1241
Db 2168 GTTGGGCGCTCCAGCAGTCAATGATGGCGGTATGCTGGCAGCAGTCAAGTCTCCAGC 2227
Qy 1242 GAATGCTCTTTCGATGGATCGGCTAAAGAAATATCGCGGTATGCTTCTCGAT 1301
Db 2228 GAGTACTTCTTCAAGGAGCGGAAGACACTGAAGACTACAGAGTATGCTCTCATCGAC 2287
Qy 1302 GCCATGACAGCAGCAGTCAAGGTAACGCCGCTACCTCCGTTAGTTCTTGAGTCT 2347
Db 2288 GCCATGACAGCAGTCAAGGTAACGCCGCTACCTCCGTTAGTTCTTGAGTCT 2347
Qy 1350 GACAAATCAAAAGTGGCGGAGGAGTGTGTGTGTGTCAGAGCAAAAGGTCAATCCAC 1409
Db 2348 GACAAAGTGTCTGCTGCTGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2407
Qy 1410 AATTTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1469
Db 2408 AAGTACATTCATATGTATGATCAATGTGTCTACAGCAGTGTGTGTGTGTGTGTGTGTGT 2467
Qy 1470 AGCTGACCCAAAGTGGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1529
Db 2468 TCTGTAGTGGAGTGTGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2527
Qy 1530 TCCCTAGCGCAGGTGGAGAGTGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1588
Db 2528 CCATCTGCCAGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2586

RESULT 5
US-08-925-230-5/c
; Sequence 5, Application US/08925230
; Patent No. 6147194
; GENERAL INFORMATION:
; APPLICANT: Coliart, Frank
; APPLICANT: Huberman, Eliezer
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,230
; FILING DATE: September 8, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,665,583
; FILING DATE: 12-AUG-1988
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..393
; OTHER INFORMATION: /note="This sequence is the top
; strand of a double stranded sequence. The sequence immedi
; ately following is the bottom strand of the double stranded sequ
; US-08-925-230-5

Query Match 17.3% Score 285.8; DB 3; Length 393;
Best Local Similarity 89.6%; Pred No. 1e-78; Indels 4; Gaps 4;
Matches 352; Conservative 0; Mismatches 37;

Qy 771 AAACAGCTGCTGTGTGGGCGGAGCCATTTGGCACTCATGAGATGACAAATATAGGCTGGA 829
Db 393 AAGCACTGCTGTGTGTGGGCGGAGCCATTTGGCACTCATGAGATGACAAATATAGGCTGGA 829
Qy 830 GTTGGCGGCGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 889
Db 335 GTTGGCGGCGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 889
Qy 890 GTTGGCGGCGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 949
Db 275 GTTGGCGGCGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 949
Qy 950 AAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1008
Db 215 AAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1008
Qy 1009 TCCGGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 1068
Db 155 TCCGGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 1068
Qy 1069 GGGCCCAAGCAGCAGTGTACAAAGTGTATGAGTATGACAGGCGCTTGGTGTCCG 1128
Db 95 GGGCCCAAGCAGCAGTGTACAAAGTGTATGAGTATGACAGGCGCTTGGTGTCCG 1128
Qy 1129 TCATTGCTGATGAGAGAAATCCAAATGTGGTGC 1161
Db 35 TCATTGCTGATGAGAGAAATCCAAATGTGGTGC 3

RESULT 6
US-08-925-230-6
; Sequence 6, Application US/08925230
; Patent No. 6147194
; GENERAL INFORMATION:
; APPLICANT: Coliart, Frank
; APPLICANT: Huberman, Eliezer
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible

```





```

: Patent No. 6147194
:
: GENERAL INFORMATION:
: APPLICANT: Collart, Frank
: APPLICANT: Huberman, Eliezer
: TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH
: TITLE OF INVENTION: AND GMP PRODUCTION
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/925,230
: FILING DATE: September 8, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 5,665,583
: FILING DATE: 12-AUG-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Wilson, Mark B.
: REGISTRATION NUMBER: 37,259
: REFERENCE/DOCKET NUMBER: ARCD:274
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (512) 474-7577
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 344 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1..344
: OTHER INFORMATION: /note="this sequence is the bottom
: OTHER INFORMATION: strand of a double stranded sequence, read from 5' to 3'. Th
: OTHER INFORMATION: top strand with which it belongs is the sequence immediately
: OTHER INFORMATION: preceding it in this listing."
:
: US-08-925-230-4

```

```

Query Match          16.7%; Score 275.8; DB 3; Length 344;
Best Local Similarity 87.8%; Pred. No. 1.2e-75;
Matches 301; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 412 GCGCCAGAGATCGCGTGGGATGTTTGGAGCCAGGCCCGCGATGTTTCTGCGGTA 471
    |||||
DB 343 GCGCCAGAGATCGTGTACCGATGTTTGGAGCCAGGCCAGCGATGTTCTGTGTA 284
    |||||
QY 472 TCCCAATCAGACGACGCGGATGGGAGCGGCTTGGGGCATCTCTCTCCAGG 531
    |||||
DB 283 TCCCAATCAGATACAGGCGGATGGGAGTGGATGGGCGATCTCTCTCAAGG 224
    |||||
QY 532 ACATTGATTTTCAAGAAGAGACATGACTGTTTCTTGAAGAGTAATGACAAGA 591
    |||||
DB 223 ACATTGATTTCTCAAGAGAGAGACATGACGCGTTCTTGAAGAGTATGACTAAGA 164
    |||||
QY 592 GGAAGACTTGGTGTAGCGCCCGGACAGATCATGACTGAAGGAGGAAATGAATTCTGC 651
    |||||
DB 163 GGAAGACTTGGTGTAGCGCCCGGACAGTCTCTGAAAGGAGAAATGAATTCTGC 104
    |||||
QY 652 AGCGACAGAAAGAAAGTTGCCATTTGTAATGAAGATGATGATTTGTGCGCATCA 711
    |||||
DB 103 AGCGAAGTAAAGAAAGTTGCCATTTGTAATGAAGATGATGATTTGTGCGCATCA 44
    |||||
QY 712 TTGCGCGACAGACTGAAGAGATGGGACTACCCACTAGC 754

```

```

DB 43 TTGCGCGACAGACTGAAGAGATGGGACTACCCCTGCGC 1
|||||
RESULT 9
US-08-629-600-1
: Sequence 1, Application US/08629600
: Patent No. 5783196
:
: GENERAL INFORMATION:
: APPLICANT: NORIEGA, Fernando
: APPLICANT: LEVINE, Myron M.
: TITLE OF INVENTION: GUA MUTANTS OF SHIGELLA
: TITLE OF INVENTION: AND VACCINES CONTAINING THE SAME
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
: CITY: Washington, D.C.
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20037
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/629,600
: FILING DATE: 9-Apr-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: KIT, Gordon
: REGISTRATION NUMBER: 30,764
: REFERENCE/DOCKET NUMBER: A-6765
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 293-7060
: TELEFAX: (202) 293-7860
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3531 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
:
: US-08-629-600-1

```

```

Query Match          13.7%; Score 226.2; DB 1; Length 3531;
Best Local Similarity 49.5%; Pred. No. 1.1e-59;
Matches 650; Conservative 0; Mismatches 648; Indels 15; Gaps 2;

QY 128 AGAGGCGTCACATACATGTTCTCATTTCTCCCTGGGTGATGAGTACTGACGAGA 187
    |||||
DB 328 AGAGGCGTCAGATTTGAGAGAGTCTCTCTGTTCTGCTGACACTACGTTCTGCCAA 387
    |||||
QY 188 CCGAGTGGACCTGACTTGTGCTGTGACCAAGAAATCACTTTAAGACCCACTGTTTC 247
    |||||
DB 388 TACTGCTGACCTAGCACCAGCGTGAAGAACTATTGCTGGAATATCCCTATGCTTTC 447
    |||||
QY 248 CTCTCCATGAGACAGTACAGAGAGCTGGATGGCATAGCAATAGCGCTTACAGCGCG 307
    |||||
DB 448 CCGAGCATATGATACGCTAAGCGAAGCGCGCTGCTATTGCTGTGGCTCAGAGAGCGG 507
    |||||
QY 308 TATGCGCTTATTCACAAAACATGTCATTGACGCGCAGAGAGAAAGATTCCCGCTGT 567
    |||||
DB 508 TATGCGCTTATTCACAAAACATGTCATTGACGCGCAGAGAGAAAGATTCCCGCTGT 567
    |||||
QY 368 GAAGAAATATGAACAGAGATTTCATCAGACAGCCGTGTGTCTCAGCGCCCAAGATGGCT 427
    |||||
DB 568 GAAGAAATATGAATCTGTGTGTGTGACTGATTCGCGACTGTTTCTGCAACGACGCT 627
    |||||
QY 428 GCGGATGTTTGGAGCCAGGCGCGGATGTTTCTGCGGTATTCCAATGACAGACAC 487

```

Db 628 GCGCAAGTGAAGAACTGACCGAGCGTAAACGTTTGGGGCTATCCGCTGTACCA 687  
Oy 488 AGCCGGATGGGAGCGCGTTGGTGGCATCATTCCTCCAGGGACATGATTTCTCAA 547  
Db 688 A-----GAAACGAACTGGTGGTATATACCGGTGTGTGCTGTTTGTAC 738  
Oy 548 AAGAGGAACTGACTGTTCTTGGAGAGATTAATGACAAAGAGGAGACTTGTGT 607  
Db 739 GCACCTGAACCGACCGCTTACGTTTACATGACCGGAAAGAGCGTCTGTACCGTGG 798  
Oy 608 AGCCCCCGGACATCAGTGAAGAGCAAAATGTAATTTGACGCGACGACGAAGGG 667  
Db 799 TGAAGGTGAAGCCCGTGAAGTGTGTGCAAA-----ATGCAAGAAACGCGTTGA 852  
Oy 668 AAGTTGGCCATTGTAATGAAGATGATGAGCTTGGCCATCATTCGCCGACAGCCT 727  
Db 853 AAAAGCGTGGTGGTGGTGAATGACGAATTCACCTGATCGCATGATACCGTGAAGACTT 912  
Oy 728 GAAGAAGATCGGAGTACCTACCTCCAAAGATGCAAGAAACAGCTGTGTGTG 787  
Db 913 CCAGAAAGGGAAGTAAACCGAACGCTGTAAAGACGAGCAAGCCGCTGTGTGTG 972  
Oy 788 GCGAGCCATTGCTCATGAGAGTACAAATAGCTGTGACTTGTGCGCCAGCTGG 847  
Db 973 TCGAGCGTGGGAGGAGGTGCGGTTAACGAGAGCGTGTGAGCGCTGTGCGCAGG 1032  
Oy 848 TGTGATGTAGTGTGTTTGGACTCTCCAGGAAATTCATCTTCCATCATCAATATGAT 907  
Db 1033 CGTGTGACGTTGTCTATGATGACTCTCCAGCGTCACTCAGAAGGTGTACTGCAAGTAT 1092  
Oy 908 CAATGATCATMAAGCAAAATACCTATCTCCAAATGATGAGGCAATGTGTGCTAC 967  
Db 1093 CCCTGAACCCCGGTAAATATCCGATGTGCAAAATATGCGCGCAACGTGCAACAGC 1152  
Oy 968 TGGCCAGGCCAAGAACCTCATTTGATGACGTGTGATGCCCTGCGGGTGGGATGGGAG 1027  
Db 1153 TCGAGGTGACCGCTCTGTGCAAGAACGTGTGTGCACTGCGGTTAAAGTGGCAATGGCCC 1212  
Oy 1028 TGGCTCATCTGATTTATCCAGAAAGTGTGCGCTGTGGCGGCCCAAGCAACACAGT 1087  
Db 1213 TGGCTTATCTGTACAACTGTATGCTGATGCGCTGCGGTGCCGACATTTACCGCTGT 1272  
Oy 1088 GTACAAGGTATGATGATGACGCGCTTGTGTGCTTCCGCTCATTTGATGAGGAAT 1147  
Db 1273 TGGTGAAGAGTAAAGCCCTGGAAGGCAACCGTATTCGGTATGCGGTATGCGGTAT 1332  
Oy 1148 CCAAAATGTGGGTATTTGCGAAAGCCTTGGCCCTTGGGGCCTCCACAGTCAATGATGG 1207  
Db 1333 TGGCTTCTCCGGGACATGCGCAAAAGCTATGCGCTGGGCGCAAGCGGTATGATAGG 1392  
Oy 1208 CTCTCTCTGCTGCGCACACACTGAGGCCCTGTGTGAATCTTTTTCGATGGATCCG 1267  
Db 1393 TTCTCATGTGGGGGTACTGGAANAATCTCGGGGTGAATGAACTTACCAAGGCCGTTC 1452  
Oy 1268 GCTAAGAAATATGCGGATGTGGTCTCTGATGCCATGAGCAAGCACTCAGACCCA 1327  
Db 1453 TTACAATCTTACCGTGTATGGTTCCTGTGGGCGGATGTCCAAAGGTTCCTCTGACCG 1512  
Oy 1328 GAACACATTTTTCAGTGAAGCTGCAAAATCAAGTGGGCCAGGAGTGTCTGTGCTGT 1387  
Db 1513 TTATTTTCCAGAGGATTAACCTTCCGCAAAACTGTGTGCGGAAGGTATGAAAGTGGCGT 1572  
Oy 1388 GCAGACAAAGGTCAATCCACAATTTTTCCTTACCTTATGCTGTGATCCATCC 1440  
Db 1573 AGCCTTAAAGTGTGCGCTGGAAGAGATCATTCACAGCAGATGTGGCGGCTGC 1625

RESULT 10  
US-09-076-761-1  
; Sequence 1, Application us/09076761  
; Patent No. 6190669  
; GENERAL INFORMATION:

APPLICANT: NORIEGA, Fernando  
APPLICANT: SZEIN, Marcelo B.  
APPLICANT: LEVINE, Myron M.  
TITLE OF INVENTION: ATTENUATED MUTANTS OF SALMONELLA  
TITLE OF INVENTION: WHICH CONSTITUTIVELY EXPRESS THE  
TITLE OF INVENTION: VI ANTIGEN  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGRUE, MIION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
CITY: Washington, D.C.  
STATE: D.C., U.S.A.  
COUNTRY: U.S.A.  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,761  
FILING DATE: 13-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: KIT, Gordon  
REGISTRATION NUMBER: 30,764  
REFERENCE/DOCKET NUMBER: A-7140  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 293-7060  
TELEFAX: (202) 293-7860  
INFORMATION FOR SEQ. ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3531 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-076-761-1  
Query Match 13.7% Score 226.2; DB 4; Length 3531;  
Best Local Similarity 49.5% Pred. No. 1.1e-59;  
Matches 650; Conservative 0; Mismatches 648; Indels 15; Gaps 2;  
Oy 128 AGACGCGCTCACCTAATAGATTTCTCATTCCTCGGGATACATGACTTCATGACAGA 187  
Db 328 AGAGCTCTGAGCTTTGACGAGCTTCTCTGCTGCTCACTACCGTTCTGGCGAA 387  
Oy 188 CCAAGTGAACCTGACTTCTGCTCTGACCAAGAAATCACTTTAAGACCCCACTGTTTC 247  
Db 388 TACTGCTGACCTCAGACCCAGCTGACGAAACTATTCCTGGAATATCCCTATGCTTTC 447  
Oy 248 CTCTCCAGGACACAGTACAGAGCGTGGGATGGCATAGCAATGGCGCTTAACAGCGG 307  
Db 448 CAGCAATGATGATACCGTAAACGGAAGCGCGCTGCTATTTGCTGTGGCTCAGAAAGCGG 507  
Oy 448 GCTGCCATTCACACACACATGTAACCTGAATTCAGGCCAATGAATTCGGAAGT 367  
Db 508 TATGCGCTTATTCACAAACATGTCATTTGAAGCCAGGAGGAAGATTTGCGCGTGT 567  
Oy 368 GAAGAAATATGAACAGGATTCATCAACAGACCTGTGTCTCAGAGCCCAAGAGTGGCT 427  
Db 568 GAAAGAAACGAAATCGTGTGTGTGATGATCCGACAGCTGTTCGCAACACAGACGCT 627  
Oy 428 GCGGATGTTTTTGAAGCGCAAGCGCGGATGTTCTCGCGGATGCCAATACAGACAC 487  
Db 628 GCGGAGTGAAGAAAGACTGACGAGGTAACGAGTTTGGGCGTATCCGCTGTTACGA 687  
Oy 488 AGCCCGATGGGAGCGCGTGGTGGGATTCATCTCTCCAGGACATGATTTTCTCAA 547  
Db 688 A-----GAAACGAACTGGTGGTATTTATCACCGGTGCGAGCTGCTTTTGTAC 738

548 AGAGGAGGAACATGACTGTTTCTTGGAGAGATTAATGACAAAGGGA  
DB 739 CGACTGTAACCAAGCCGGTTAGCGTTTACATGACGCCGAAGACGCTCT  
QY 608 AGCCCCCGGACCATCACACTGAAGGAGCAAAATGTAATTTGACAGCG  
DB 799 TGAAGGTGAAGCCCGTGAAGTGTGCTGGCAAA-----AATGACAGA  
QY 668 AAGTGGCCCATTTGAATGAAGATGATGAGCTTGTGGCATCTTGGC  
DB 853 AAAAGCGCTGGTGGTTGATGACGAATTCACCTGATGGCATGATCAG  
QY 728 GAAGAAGATCGGAGCTACCCACTGACCTCCAAAGATGCCAAGAAACA  
DB 913 CCAGAAAGCGGAAGCTAAACCGACCGCTGTAAGACGAGCAAGCGCG  
QY 788 GGCAGCCATTGGCACTCATGAGAGATGACAGATATAGGCTGGACTTGT  
DB 973 TGCACCGGTTGGCGAGGTTGGCGGTAACGAGAGCGTGTGGACGCGCT  
QY 848 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 1033 CGTTGACGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 908 CAAGTACATCAAGCAAAATACCTTAATCTCCAAGTCAATTGGAGGCA  
DB 1093 CCGTGAACCCGCTGCTAAATATCGGATCTGCAAAATTAACGCGGCA  
QY 968 TCCCGAGCGCAAGCAACCTATGATGACAGGTGATGACCTGCGCGGT  
DB 1153 TGCAGGTGACCGCTGCTGCGCAAGAGCTGTTGCAAGTGGCTTAAGT  
QY 1028 TGGCTCCATCTGATTAATTCAGGAAGTGTGCTGTGGCGCGCCCA  
DB 1213 TGGCTCTATCTGTACACTGCTATGCTGCTGCTGCTGCTGCTGCTGCT  
QY 1088 GTACAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 1273 TGTGACGACAGTAAGAGCCCTGGAAGGACCGGTATTCGGTATTCG  
QY 1148 CCAAAATGTGGTCAATATTCGCAAGCCCTTGGCCCTTGGCGCTCCAC  
DB 1333 TCGCTTTCGCGGACATCGCCAAAGCTATCGCCCTGGCGCAAGCGC  
QY 1208 CTCTCTCCCTGCTGACCACTAGAGCCCTGTGATTAATCTTCTTTC  
DB 1393 TTCCATGCTGGGGGTACTGGAAGATCTCGGGTGAATTCGAACTCTA  
QY 1268 GCTAAGAATAATGCGCGTATGCTTCTGTGATGCCATGACAAAGCA  
DB 1453 TTACAAATCTTACCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 1328 GACAGATATTTCAGTGAAGCTGACAAATCAAGTGGCGCCGAGGT  
DB 1513 TTATTTCCAGAGCGATTAACCGTCCGACAAACTGCTGCGGAGGTAT  
QY 1388 GCAGGACAAAGGCTAATCCAAATTTGCTCCCTTACCTGATGTTGGT  
DB 1573 AGCCTATTAAGGTCCCTGAAAGAGATCATTCACGACGAGATGGCG  
RESULTS  
US-09-134-001C-2187  
Sequence 2187, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucelle-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2187  
LENGTH: 1485  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2187  
Query Match 10.8%; Score 179.4; DB 4; Length 1485;  
Best Local Similarity 47.5%; Pred. No. 2.1e-45;  
Matches 642; Conservative 0; Mismatches 696; Indels 15; Gaps 3;  
QY 136 TCACCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 50 TAAACATTCAGCAGCGTGTACTCATTCAGCTGATCAGATGATGATGATGAT  
QY 196 ACCGTACTTCTGCTGACCAAGAAATCCTTTAGACCCCACTGTTCTCTCCCA  
DB 110 ACTTAAGTCAAAATTAATCAGATGATGATGATGATGATGATGATGATGAT  
QY 256 TGGACACACTCAGAGAGGTGGGATGGCCATGACATGCGCTTACAGCGGTATTGCT  
DB 170 TGGATACAGTAACTGAATCAAAATGCAATGCTGATGCTGACAGAGCGGTTAAGGTG  
QY 316 TCATCCACCACACACTGATCACCCTGAATTCAGGCCAATGAAAGTGGAAAGTGAAGAAAT  
DB 230 TTATTCATTAAGAAATATGAGCGCTGCAAGAGCAAGCTGATGATGATGATGATGAT  
QY 376 ATGACAGGAGATTCATCAGACACCTGTGCTCTCAAGCCCAAGAGATGCGCTGCGGATG  
DB 290 CAGAAATGCTTATTTCTTAACCGCTTCTTCTTAACACCGGAAAGTGTATGAGG  
QY 436 TTTTGGAGCCCAAGGCCCGGCGATGTTCTGCGTATGCCAATTCACAGACAGCCCGGA  
DB 350 CTGAAGCATTAATGGGTAAATACCGTATCTGCTGCTGCTGCTGCTGCTGCTGCT  
QY 496 TGGGAGCGCGTGGTGGGATCATCTCTCCAGGACATGATTTTCTCAAGAGAGAGG  
DB 407 AGGATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 556 AACATGACTGTTTCTTGAAGAGATTAATGACAAAGAGGAAAGTGTGATGAGCCGCC  
DB 467 CAATTAATAATATCAATGATTAATGAC-----GAAGATTAATTAATACAGCTCCAG  
QY 616 GCAGCATCACTGAGAGAGCAAAATGAAATTTCTGACGCGACGACAGAAAGGAAGTTGG  
DB 518 TTGTTACGACTTATGATGAGCGGAGGCTATTTCAAAACATTAAGATTGAGAACTTTC  
QY 676 CCATTGTAATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 578 CATTTAGTA---GAAATGCTGCTTTAGAGAGATTAATCACTATTAAAGATATTGAAAAAG  
QY 736 ATGCGGACTACCCACTAGCCTTCAAGATGCCAAGAAACAGCTGCTGTGGGAGAGCCA  
DB 635 TACTTGAATTTCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 796 TTGGCAGTCATGAGATGACAAAGTATGATGATGATGATGATGATGATGATGATGATGAT  
DB 695 TCGGTACGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 856 TAGTGGTTTGGACTTTCGCCAGGAAATTCATCTTCAGATCAATATGATCAAGTACA  
DB 755 CATTAATTAATGATTAACACTCATGCTCATTTGAAAGGGGTTATTAATCAAGTTAAACACA  
QY 916 TCAAGAGCAAAATACCTTATCTCCCAAGTCAATGGAGCAATGGTGTGCTGCTGCGCAGG  
DB 815 TCAAGAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT  
QY 976 CCAAGAACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
DB 875 CACGTGCTTTATTTGAAGCGGCTGCGATGTTGTTAAGTATGATGATGATGATGATGATGAT

QY	1036	TCGTCATTATCCAGAGAGTCGTGGCCCTGTGGGGGGGGCCCGAACAGACAGCAGTGTCCAAAG	1095
Db	935	TTTCCACAAACACCGTGTGTTGTCACAGGCTGTAGCTGTGCCCTAAATTACAGCAGTTTATGATT	994
QY	1096	TGATAGAGTATGCACGGCGCCTTGGTGTTCGGCTCATTCGTATGGAGAGATCCAAATG	1155
Db	995	GTGCTACAGAGAAGCCCCGTAAACCATGTATGAAGCTATTATTGCTATGATGGTATTAAGTCTT	1054
QY	1156	TGGGTCAATTGGCAAGCCTTGGCCCTTGGGGCCCTCCACATGTCATGATGGGCTCTGCC	1215
Db	1055	CAGGTGATATTATCAAAAGCAATTAGCTGCTGTGTCATGCGGTATGATGAGGTAGTTGTT	1114
QY	1216	TGGCTGCGACCACTGAGAGCCCTTGTTGGTAATCTCTTTTCCATGGATCCGGCTAAAGA	1275
Db	1115	TAGCTGTAGCAGAAAGAAAGTCCTGTGTSCAATCAATATTCCAAAGTATACAAATATTAAG	1174
QY	1276	AATATCGCGGTATGGGTTCTCTCGATGCCATGTGACAAAGCACTCCAGACCCAGAACAGAT	1335
Db	1175	TTTATTCCGGGCGATCGGATCTTTTAGTGTCATTATGAAAAAGTTTCAATATGATCTGTTACTTC	1234
QY	1336	ATTTCAGTAGAAGCTACAAAATCAAGTGGGCCAGGAGAGTCTGGTGCCTGTGCAGAGCA	1395
Db	1235	AAGAAAGATAAAACCCCAAGAAAATTTGTCTCTGAAGGTATGAAAGTCTGTACAGCTTTATA	1294
QY	1396	AAGGTCATATCCACAAATTTGTCCCTTACCTGATTCGTGGCATCCAAACATCATGCCAG	1455
Db	1295	AAGGCACATTACGATGATACAAATTATCAACTTATGGGTGGCGTTATAGCTGCCATGGGTT	1354
QY	1456	ACATTGGTCCAAAGAGCTGTGACCCAGTCCGAG	1488
Db	1355	ATACGTGGTTGAGAAACCTTAAAAAAATTTACGTG	1387

RESULT 12  
 US-09-221-017B-497  
 Sequence 497, Application US/09221017B  
 Patent No. 6444799  
 GENERAL INFORMATION:  
 APPLICANT: Ross, Bruce C.  
 TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
 NUMBER OF SEQUENCES: 1120  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 PAGE MILL ROAD  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows  
 SOFTWARE: FASTSEQ for Windows Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/221.017B  
 FILING DATE: 23-DEC-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PP1182  
 FILING DATE: 31-DEC-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PP1546  
 FILING DATE: 30-JAN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PP2911  
 FILING DATE: 09-APR-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/AU98/01023  
 FILING DATE: 10-DEC-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Montoy, Gladys H  
 REGISTRATION NUMBER: 32,430

```
? REFERENCE/DOCKET NUMBER: 27340-20021.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-813-5600
? TELEFAX: 650-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ ID NO: 497:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5884 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: circular
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: UNKNOWN
? ORIGINAL SOURCE:
? ORGANISM: PORPHYROMONAS GINGIVALIS
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1...5884
? US-09-221-017B-497
```

Query Match	Similarity	9.3%	Score 154.6	DB 4	Length 5884
Best Local	Similarity	49.6%	Pred. No. 2.4e-37		
Matches	499	Conservative	0	Mismatches 48	Indels 24
				Gaps	3
QY	222	ATCACTCTTTAAGACCCCACTGGTTTCTCTCTCCCATGAGCACAGTCACAGAGGCTGGATG	281		
DB	2706	ATCAAGCATGAATATATCCCGTTTCTTGGCCATCATGCAATCGGTTCCAAATGATACGGTG	2765		
QY	282	GCCATAGCAATGGCCCTTACAGCGCGTATGGCTTCATCCACACACACTGTACACCTGAA	341		
DB	2766	GCCATCGAATTTGGCAGCGCAACGGGGAGCTTTCGTTATCTTCGGATTCGACAGCATAGAG	2825		
QY	342	TTCCAGGCCCAATGAAAGTTTCGGAAGTGAAGAAATATGAACAGGATTCATCCACAGCCT	401		
DB	2826	AATCAGAGCCGAATATGGTTCTGCTGTGAAGAAATTCAGAGCCGGATTTGTACACAGCAT	2885		
QY	402	GTGTCTTCAGCCCCAAGCATGCGCTGGCGGATGTTTTTGTAGGCCCAAGGCCCGCATGCT	461		
DB	2886	TCCATCTCAGACCCGCAATACGCTGGCGCGATTTGGATCTGTGTCMAAAGCAGGA	2945		
QY	462	TTCTCGGGATCCCAATCACAGCACAGGCCGATGAGGGAGCGCGTGGTGGGCAATCATC	521		
DB	2946	CATTAATTAATATGATATCACTACACAGCGTTTCTCCATATGGCCGCTGTGATGGGTATCGTC	3005		
QY	522	TTCCTCCAGGACATTGATTTTCTCAAAGAGAGAAATGATGATGTTTTTGGGAAGATA	581		
DB	3006	ACGACAGAGAGAC---TATCGCTCAGCACAGCACTCACCTCGAAGCGGTCAAAAGTTTTT	3062		
QY	582	ATGACAAAGAGGGAAGACTTGGTGGTAGCCGCCCGCAGCATCACACTGAAGAGGCCAAT	641		
DB	3063	ATGAGCGCTTTCGAGAGACTTACGGTGGCGCAGTGGGTATTAACGCTAGGCAAGCCAAAC	3122		
QY	642	GAAATTTCTGACCGCAGCAGAAAGGAAGTTGCCATTGTAAATGAATGATGACTT	701		
DB	3123	GATATTATTATTTGGGAAACCAACTCAATACGCTGCCGATATCCAGCAAAATGAGATTCG	3182		
QY	702	GTGGCCATCATTTGCCCGCAGCACTGAAGAAATTCGGGATTCACCATAGCTTCCAA	761		
DB	3183	CAGTATTTTCGTTTCCGTTAAGATTATACGACAGCCACAGAACAAATTCGTCGAACCTTCG	3242		
QY	762	GATGCCAAGAAACAGCTGCTGTGTGGGGCAGCCATTGGCATTATGAGAGATGACAGTAT	821		
DB	3243	AACCTACAGCAGCAAGACCCCTCTCGTGGGTGCGGATATCAATACAGGCACTATAAGAG	3302		
QY	822	AAGCTGGACTTCTCTGCCCCAGGCTGTGTGATGATGTGTTTTGGACTCTTCCAGGA	881		
DB	3303	CGAGTACCTGCACCTGGTGGAGCCGGTGTGGATGTGGTGTGCATGCATCATCGGACGCG	3362		
QY	882	AATTCGATTTCCAGATCAATATGATCAAGTCAATCAAAAGACAAATATC---CCTATATC	938		
DB	3363	TATTCGAGATGGCAGCATACGATCCCTTGATATCAAGAAATATCGGAGCACTTTG	3422		



Query Match	4.9%	Score 81.6	DB 1	Length 1843
Best Local Similarity	46.9%	Pred. No. 5e-15	Mismatches 255	Conservative 0
772	ACACAGCTGCTGTGTGGGGCAGCCATTGGCAGCTACATGAGGATGACAGATATAGGCTGACT	831		
Db	625	AGCATCTGGCTGCCACCTCAGCAGCAGGCTCTTCTGACTTTGAGCAGCGATGGAAAGATCC	684	
QY	832	TGCTCGCCCAAGGCTGTGTGGATGTAGTGGTTTGGACTCTTCCACAGGAATTCATCT	891	
Db	685	TGGAAGCTATTCCCAAGGTAGTATATATAGCTGGATGTGGCAATGGCTACTTGGAAC	744	
QY	892	TCCAGATCATATGATCATCAATACATCAAGSACAAATACCCCTATCTCCCAAGCATTTGAG	951	
Db	745	ACTTGTGTAATTTGTAAAGATGTACGGAAAGCGCTTCCCGACACACCATCATATGGCAG	804	
QY	952	GCAATGTGTGACACTGCTGCCAGGCCAAGAACCTCATTTGATGAGGTGTGATGCCCTGC	1011	
Db	805	GGAATGTGTGTAACAGAGAGAGATGTAGAAAGCTCATTCCTTCTGGGCGTGAACATCATCA	864	
QY	1012	GGGTGGGCATGGGAAGTGGCTCCATCTGCATTTATCCAGGAAGTGTGCTGGCTGGGGCGC	1071	
Db	865	AAGTGGGAATTTGGCGCAGGCTCTGTGTGTACTGTCTGGAAAGAACTGGAGTGGGGTATC	924	
QY	1072	CCCAAGCAGACAGCGTGTACAGGTGTATAGATGACAGCGGCTTGGTGTCCGCTCA	1131	
Db	925	CACAGCTCAGCGAGTGTAGAGTGTGCAATCTGCTCATGGCCTCAAAAGCCACATCA	984	
QY	1132	TTTGGTATGAGAGAAATCCAAATGTGGGTATATTGGAAAGCCTTGGCCCTTGGGGCCT	1191	
Db	985	TTTCAGATGAGAGTTTGCAGCTGTCTGTGGGATGTGGGCAAGGCTTTTGGGGCAGSAGCTG	1044	
QY	1192	CCACAGTCATGATGGGCTCTCTGCTGGCTGCCACACTGAGGCCCTGGTGAATACTTCT	1251	
Db	1045	ACTTCGTGATCTGTGGTGGCTGCTGCTGGCGACACTGTGATGAGTGGAGCCTCATCG	1104	
QY	1252	TTTCCGATGAGGATCCGGTAAAGAAATATGGGGTATGGGTCTCTCGATGCCATGACA	1311	
Db	1105	AGAGGAGTGGCAGAGATGACACAGCTCTTATGGAAATGAGTTCTGAATGGCATGAAGA	1164	
QY	1312	AGCA 1315		
Db	1165	AGTA 1168		

RESULT 15

US-08-961-527-126/C

Sequence 126, Application US/08961527

Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATE:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

```

NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 126:
SEQUENCE CHARACTERISTICS:
LENGTH: 13121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-126

Query Match          4.6%: Score 75.8; DB 4: Length 13121:
Best Local Similarity 47.7%: Pred. No. 1,1e-12:
Matches 255; Conservative 0; Mismatches 277; Indels 3; Gaps 1.

QY      767  CAAGAACAAGCTGCTGCTGTGTGGGCGACCATTTGGCAGTCATGAGATGACAGTATAGCT  826
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12794 CAATCAAGAGGCTCATTTGCTTCTATCTCTGTGCGTGAATTAAGGATTAAGTATAGTAAATTTGCT 12735

QY      827  GACACTTGCTCGCCCAAGCGCTGCTGTGTGATGTCTAGTGGTTTGGACTCTCCACAGGAATTC  886
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12734 TAGCCAGCTCAAGCGTGTATACTCTGTAGTATATCACTATTGTATTTGCTCAATGGTTCATTC 12675

QY      887  CATCTTCCAGCATCATATGATCATCACTACATCAAGAACAATACCCCTTAATCTCAAGTCAT  946
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12674 GGAATGACCGGATTTCTATATCATTCACACACTCAAGAAAGAAATTGCCAGATACCTTTGTCTAT 12615

QY      947  TGGAGGCAATGTGTGTACTGCTGTGCCAGGCCCAAGAACCCTCATTTGATGACAGGTGTGGATGC  1006
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12614 TCGTGAATAATGTGGGCAACCCAGAGAGCTGTGCGTGAATTGGAAAAAGTGTGGCGGATATC 12555

QY      1007  CCTGGGGGTGGGCATATGGGAGAGTGGGCTCCATTCGATTATCCAGGAAGTGCCTGGCTGTGG  1066
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12554 TACTAAGGTGGAATTCGGTCTGTGTAGTGTGTATACCAAGGTTAAGACATGTTTGG 12495

QY      1067  GCGGCCCCCAAGCAACAGCAGTGTACAAAGGTATAGTATGCACGGCGCTTTGGTGTTC  1126
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12494 TA---CAGGTGGTGTGGCAGTGTGGCTGCTCTACGCTGTGTGGCCAAAGGCTGCGGTAAACC 12438

QY      1127  GGTGATTTGCTGATGAGGAAATCCAAATGTGGGTCAATTTGCCAAAGCCCTGGCCCTTTGG  1186
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12437 GATTATACGCTGTATGGAGGAATTCGTACTACAGGCTGATATTGCTATATCCGCTTCCG 12378

QY      1187  GGGCTCCACAGTCATATATAGGGCTCTCTCCGTGGCTGCACACACTAGAGCCCTGTGTGAATA  1246
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12377 TGTCTACATGATCATATATATGTGTTCCCTCTTTTGGACGACACATTTGAAGTCCAGAGAAAC 12318

QY      1247  CTTCCTTTTCCGATGGGATCCGGCTAAAGAAATATTCGGCGTATGGGTTCTTCGAT  1301
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 12317 GATTGAAGTCGATGCTGAGAACAGTTCAAGAAATATTATGTGTTACGCTCCACAAATAT 12653

```



Greene, Mark Keltman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: RMAL Plate: 3 Row: J Column: 13  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 1702963  
 This clone has the following problem: frame shifted.

## FEATURES

source

1. 1691

/organism="Homo sapiens"

/db\_xref="LocusID:3615"

/db\_xref="taxon:9606"

/clone="IMAGE:2988188"

/tissue\_type="Colon, adenocarcinoma"

/clone\_lib="NTL\_MGC\_15"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

BASE COUNT 467 a 411 c 448 g 365 t

ORIGIN

Query Match 96.8%; Score 1601; DB 11; Length 1691;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1618; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 17 CGAGACACGCGCGGTCTCTGTGTGGCCATGGCCGACTACTGATTAGTGGGGCAC 76  
 DB 10 CGAGACACGCGCGGTCTCTGTGTGGCCATGGCCGACTACTGATTAGTGGGGCAC 69  
 QY 77 GTCTACGTCACAGACGAGCTACACACAGCAGCTCTTCACTGGGAGAGCGCTT 136  
 DB 70 GTCTACGTCACAGACGAGCTACACACAGCAGCTCTTCACTGGGAGAGCGCTT 129  
 QY 137 CACCTACATGACTTCTCTATCTCCCTGGTACATGACTTCACTGACAGACAGTGA 196  
 DB 130 CACCTACATGACTTCTCTATCTCCCTGGTACATGACTTCACTGACAGACAGTGA 189  
 QY 197 CCTGACTTCTCTGACCAAGAAATCACTTTAGACCCCACTGTTCTCTCCAT 256  
 DB 190 CCTGACTTCTCTGACCAAGAAATCACTTTAGACCCCACTGTTCTCTCCAT 249  
 QY 257 GCACACAGTCACAGAGGCTGGGATGCCCATAGCAATGGCGCTTACAGCGGATTTGCTT 316  
 DB 250 GCACACAGTCACAGAGGCTGGGATGCCCATAGCAATGGCGCTTACAGCGGATTTGCTT 309  
 QY 312 CATTCCACCACTGACCTGATTCAGAGGCAATGAAGTTGGAAAGTGAAGAATA 376  
 DB 310 CATTCCACCACTGACCTGATTCAGAGGCAATGAAGTTGGAAAGTGAAGAATA 369  
 QY 377 TGAACAGGGATTCATCACAGACCCCTGTGCTCAGCCCCCAAGATCGCGTGGGATGT 436  
 DB 370 TGAACAGGGATTCATCACAGACCCCTGTGCTCAGCCCCCAAGATCGCGTGGGATGT 429  
 QY 437 TTTTGAAGCCCAAGGCCCGCATGTGTTCTGCGGTATCCCAATCACAGACAGCGGAT 496  
 DB 430 TTTTGAAGCCCAAGGCCCGCATGTGTTCTGCGGTATCCCAATCACAGACAGCGGAT 489  
 QY 497 GGGAGACCCCTGTGGTGGGATCATCTCTCCAGGGCATTGATTTCTCAAGAGGAGGA 556  
 DB 490 GGGAGACCCCTGTGGTGGGATCATCTCTCCAGGGCATTGATTTCTCAAGAGGAGGA 549  
 QY 557 ACATGACTGTTTCTTGAAGAGTAAATGACAAAGAGGAAGATTGGTGGTGGCCCGG 616  
 DB 550 ACATGACTGTTTCTTGAAGAGTAAATGACAAAGAGGAAGATTGGTGGTGGCCCGG 609  
 QY 617 CAGCATCACATGAGAGGAGCAAAATTTGACGCGCAGCAGACAGAGGAAAGTTGGC 676  
 DB 610 AAGCATCACATGAGAGGAGCAAAATTTGACGCGCAGCAGACAGAGGAAAGTTGGC 669  
 QY 677 CATTGTAATGAAGATGATGAGCTGTGGCCATCATTTGGCCGACAGACCTGAAGAAGA 736  
 DB 670 CATTGTAATGAAGATGATGAGCTGTGGCCATCATTTGGCCGACAGACCTGAAGAAGA 729

QY 737 TCGGGACTACCCACTAGCCTCCAAAGATGCCAAGAAACAGCTGCTGTGGGGCAGCCAT 796  
 DB 730 TCGGGACTACCCACTAGCCTCCAAAGATGCCAAGAAACAGCTGCTGTGGGGCAGCCAT 789  
 QY 797 TGGCACTATGAGGATGACAAATATAGGTGGACTTGTCTGCGCCAGGCTGTGTGATGT 856  
 DB 790 TGGCACTATGAGGATGACAAATATAGGTGGACTTGTCTGCGCCAGGCTGTGTGATGT 849  
 QY 857 AGTGGTTTGGACTCTTCCAGGAAATTCATCTTCCAGATCAATATGATCAATACAT 916  
 DB 850 AGTGGTTTGGACTCTTCCAGGAAATTCATCTTCCAGATCAATATGATCAATACAT 909  
 QY 917 CAAAGCAAAATACCTTAATCTCCAAGTCATTTGAGGCAATGTGTCACTGTGCCAGGC 976  
 DB 910 CAAAGCAAAATACCTTAATCTCCAAGTCATTTGAGGCAATGTGTCACTGTGCCAGGC 969  
 QY 977 CAAAGAACCTATGATGACAGTGTGATATGCTTGGGGTGGGCATATGGAGATGCTCCAT 1036  
 DB 970 CAAAGAACCTATGATGACAGTGTGATATGCTTGGGGTGGGCATATGGAGATGCTCCAT 1029  
 QY 1037 CTGCATTATCCAGAAAGTGTGCTGGCCTGTGGGGGGCCGCCCAAGCAACAGCAGTGTACAGGT 1096  
 DB 1030 CTGCATTATCCAGAAAGTGTGCTGGCCTGTGGGGGGCCGCCCAAGCAACAGCAGTGTACAGGT 1089  
 QY 1097 GATAGATATGACAGCGGCTTTGTGTTCGGTCAATGTCTGATGAGAGAAATCAAAATGT 1156  
 DB 1090 GATAGATATGACAGCGGCTTTGTGTTCGGTCAATGTCTGATGAGAGAAATCAAAATGT 1148  
 QY 1157 GGGTCATTTGGGAAAGCTTTGGCCCTTTGGGGCTTCCACAGTCATGATGGCTCTCTCT 1216  
 DB 1149 GGGTCATTTGGGAAAGCTTTGGCCCTTTGGGGCTTCCACAGTCATGATGGCTCTCTCT 1208  
 QY 1217 GCGTCGACACACTGAGAGCCCTGTGTGAATCTTTTCCATGAGGATCGCGCTAAAGAA 1276  
 DB 1209 GCGTCGACACACTGAGAGCCCTGTGTGAATCTTTTCCATGAGGATCGCGCTAAAGAA 1268  
 QY 1277 ATATCGGGTATGGGTTCTCTGATGCCATGAGCAACAGCAGCAGCAGAGAGATA 1336  
 DB 1269 ATATCGGGTATGGGTTCTCTGATGCCATGAGCAACAGCAGCAGCAGAGAGATA 1328  
 QY 1337 TTTCACTGAAGCTGACAAATCAAAATGGGCCAGAGAGTGTCTGTGCTGTGACAGACA 1396  
 DB 1329 TTTCACTGAAGCTGACAAATCAAAATGGGCCAGAGAGTGTCTGTGCTGTGACAGACA 1388  
 QY 1397 AGGTCATTCACAAATTTGTCCCTTACTGATGCTGTGCATCAACACTCATATGACAGA 1456  
 DB 1389 AGGTCATTCACAAATTTGTCCCTTACTGATGCTGTGCATCAACACTCATATGACAGA 1448  
 QY 1457 CATTGTGCCAAGAGCTTACCCAAAGTCCAGCCATGATGTACTGTGGGAGCTTAAGTT 1516  
 DB 1449 CATTGTGCCAAGAGCTTACCCAAAGTCCAGCCATGATGTACTGTGGGAGCTTAAGTT 1508  
 QY 1517 TGAGAAAGAAAGCTCTCTACGCCAGGTGGAAGTGGCGCTCATAGCTTCATTCGATGA 1576  
 DB 1509 TGAGAAAGAAAGCTCTCTACGCCAGGTGGAAGTGGCGCTCATAGCTTCATTCGATGA 1568  
 QY 1577 GAAAGGCTTTTCTGAAAGGATCCAGCACACCTCTGGTTTTTTTTCATAAAAGT 1636  
 DB 1569 GAAAGGCTTTTCTGAAAGGATCCAGCACACCTCTGGTTTTTTTTCATAAAAGT 1628  
 QY 1637 TTAGAAAAGA 1645  
 DB 1629 TTAGAAAAGA 1637

RESULT 2  
 AK002676 1655 bp mRNA linear HTC 19-JAN-2002  
 LOCUS AK002676  
 DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610025012:inosine 5'-phosphate dehydrogenase 2, full insert sequence.  
 ACCESSION AK002676  
 VERSION AK002676.1 GI:12832835



**KEYWORDS**  
HTC: CAP trapper.  
**SOURCE**  
Mus musculus (strain:C57BL/6J) adult male kidney cDNA library  
clone:0610025012.

**ORGANISM**  
Mus musculus  
Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

**REFERENCE**  
1 Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
PUBMED 10349636

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sakaki,N., Carninci,P.,  
Konno,H., Akiyama,Y., Nishi,K., Kitsuai,T., Tshino,H., Itoh,M.,  
Sunji,N., Ishii,Y., Nakamura,S., Hazama,N., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaiguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohira,E., Wachiaki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kakai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
RIKEN Integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
Aizawa,T., Hara,A., Fukunishi,Y., Konno,H., Machi,J., Fukuda,S.,  
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kodoko,S., Yamanaoka,I.,  
Saito,T., Okazaki,Y., Gojodori,T., Bono,H., Kakuwa,T., Saito,R.,  
Kadoita,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiya,H.,  
Kuehl,P., Lewis,S., Matsuo,Y., Ninkaido,I., Pestle,G.,  
Quackenbush,J., Schriml,L.M., Staudt,F., Suzuki,H., Tomita,M.,  
Wagner,E., Mashio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
Madorelli,R., Barsh,G., Blake,J., Boftelli,D., Botunga,N.,  
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bulc,C.,  
Fletcher,C., Fujita,M., Galaboldi,M., Gustinchin,S., Hill,D.,  
Hotmann,M., Hume,D.A., Kamlaya,M., Lee,N.H., Lyons,P.,  
Marchionni,L., Mashima,J., Mazzarelli,J., Momcherts,P., Nordone,P.,  
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
Sato,R., Schonbach,C., Seya,T., Shibata,Y., Struch,K.F., Suzuki,H.,  
Toyo-Oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,  
Wyshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohzuki,S.  
and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

5 (bases 1 to 1655)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bulc,C.,  
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,  
Hara,A., Hayatsu,N., Hill,D., Hitamoto,K., Hirooka,T., Horii,F.,  
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,  
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
Numasaki,R., Ono,R., M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Schriml,L., Shibata,K., Shihata,Y., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
Tanaka,T., Tejima,Y., Toya,T., Yamanura,T., Yamanashi,T.,  
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and  
Hayashizaki,Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group; RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-72 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGGACGAGGGCGCCCAACTGCATTTTCTTTTTTTTNN 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to kot = 5.0. Second strand cDNAs was prepared with the primer adapter of sequence 5'-GAAGAAGACAGAGCATCACAAGCTCAATTAAATTAATAAACCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES	SOURCE
Location/Qualifiers	1..1655
/organism="Mus musculus"	
/strain="C57Bl/6J"	
/db_xref="FANTOM_DB:0610025012"	
/db_xref="MGD:MGI:1896782"	
/db_xref="taxon:10090"	
/clone="0610025012"	
/sex="male"	
/tissue_type="kidney"	
/clone_id="RIKEN full-length enriched mouse cDNA library 1..1655"	
/gene="Impdh2"	
273..1610	
/gene="Impdh2"	
/note="data source: MGD, source key: MGI:109367, evidence: ISS"	
inosine 5'-phosphate dehydrogenase 2 putative"	
/codon_start=1	
/protein_id="BAB22278.1"	
/db_xref="GI:12832836"	
/db_xref="MGD:MGI:109367"	
/translation="MDYTEAGMA IAMLTGIGIFIHNCNTPFOANEVRKKKEQ FIDPVPVSPKRDYDEAKARHGFCIPITDRGMSRLVGIISSDIIDLPKEEE DPELEIMLTREKEDLVPAAGVTALKEANFIILPDRCGRAPINDELDAIARDLKR KNRDPASDKAKRLLICGADICTHEDDKVDLDIALAGVDVIVDSSGNSFOI NM KYIKEXPSLDIOVIGSNVYSOAKNLIDAGVDALRVNGSSICTTTOYLEAGFRQA AVTVSETRFERFYPVALDGGITONNGHIAKALAUCASTVMGSGSLTTOYLEAGFRQA DGIRLKRYRGMSLDAMDRIHSQNRFSEADKIKVAGVGAVADKDSIKHFVYLL AGDISODIDGAKSLDTORAWMYSGELTFEKRTSAOEGGVHSLHSEKRLF"	
1642..1647	
/gene="Impdh2"	
/note="putative"	
1655	
/gene="Impdh2"	
/note="putative"	
BASE COUNT	408 a 403 C 452 g 392 t
ORIGIN	
Query Match	78.9%; Score 1305.4; DB 11; Length 1655;
Best Local Similarity	88.2%; Pred. No. 0;
Matches 1431:	Conservative 0; Mismatches 191; Indels 1; Gaps 1

```

7  GGGCGGTCTTCGGAGACACGGCGGCTGTCTGTGTTGGCCATGCCCGCACTACCTGATTA 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
db  GCGCGGCTTCTGGAGACCACGGCTGTCTCTCTCCGTGCTATGGCGGCACTACCTGATTA 85
    GTGGGGGACGCTCCCTACGCTGGCCAGACGAGCACTCACAGCAGCAGCACTCTTCACTGGC 126

```

Db	1165	TCCAAAATGTGGGTATATTGGCAAAAGCTTTGGCTTTGGGGCTTCCACAGCTATGATNG	1224
Qy	1207	GCTCTCTCTGGCTGCCACCACTAGAGCCCTGGTGAATACTTCTTTCCGATGGGATCC	1286
Db	1225	GTCTTCTCTGGCTGCCACCACTAGAGCCCTGGGATGATCTTCTTCAATGGGATTC	1294
Qy	1267	GGCTAAAGAAATATGCGCGTATGGCTTCTTCGATAGCATGAGACAGCAGCTCAGAGCC	1326
Db	1285	GGCTAAGAAATATGCGAGATGATGGTCTCTTGATGGCATGAGCAACATCTCAGAGCC	1344
Qy	1327	AGACAGATATTTTCACTGTAAGCTGACAAATTCAAAGTGGCCAGGAGTGTCTGCTG	1386
Db	1345	AGAAGCGATACTTCACTGTAAGCTGACAAATTCAAAGTGGCCAGGAGTGTCTGCTG	1404
Qy	1387	TGCAGACAAAGGCGATCCCAAAATTTGTCCTTACTCATTTGTCGATCCAGCACT	1446
Db	1405	TGCAGACAAAGGCGCTCTATCCCAAGTTCCTTACTCATTTGTCGATCCAGCACT	1464
Qy	1447	CATGCCAGACATTTGGTGGCCAAAGAGCTTGACCCCAAGTCGAGGCCATGATGACTG	1506
Db	1465	CTGTGACAGACATTTGGTGGCCAAAGAGCTTGATACCCAAAGTCGAGGCCATGATG	1524
Qy	1507	ACCTTAAGTTTGAAGAAGAACGCTCTCAGCCAGAGTGAAGGTGGCTCATAGCTCC	1566
Db	1525	ACCTTAAGTTTGAAGAAGAACGCTCTCAGAGTGAAGGTGGCTCATAGCTCC	1584
Qy	1567	ATTGCGATGAGAGGGGCTTTTCTGAAAAGGATCCAGCACACCTCCTCGGTTT	1626
Db	1585	ATTGCGATGAGAGGGGCTTTTCTGAAAAGGATCCAGCACACCTCCTCGGTTT	1644
Qy	1627	CAA 1629	
Db	1645	AAA 1647	
RESULT 3			
LOCUS	AK011777		
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length		
	cDNA library, clone:2610044P12; Inosine 5'-phosphate		
	dehydrogenase 2, full insert sequence.		
ACCESSION	AK011777		
VERSION	AK011777.1	GI:12848115	
KEYWORDS	HTC; CAP trapper;		
SOURCE	Mus musculus (strain: C57BL/6J) 10 days embryo cDNA to mRNA,		
	clone:11b; RIKEN full-length enriched mouse cDNA library		
	clone:2610044P12.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	1 Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
TITLE	Kono, H., Akizawa, Y., Nishi, K., Katsunaka, T., Tashiro, H., Itoh, M.,		
JOURNAL	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
MEDLINE	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
PUBMED	Fujisawa, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watabiki, M.,		
REFERENCE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsutani, S., Kawai, J.,		
AUTHORS	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format		





|||||  
Db 477 AATCCAAATGTGGTCATATTTCCGAAGCCTTGCGCCCTTGGGCGCTCAGACTCATGAT 418  
Oy 1205 GGGCTCTCTCTGGCTGCGCCACCTAGAGCCCTGTGTAATCTTCTTTCGATGGAT 1264  
Db 417 GGGCTCTCTCTGGCTGCGCCACCTAGAGCCCTGTGTAATCTTCTTTCGATGGAT 358  
Oy 1265 CCGGCTTAAGAAATATCGGGGTATGGTCTCTCGATGCCATGGACACACCTCAGCAG 1324  
Db 357 CCGGCTTAAGAAATATCGGGGTATGGTCTCTCGATGCCATGGACACACCTCAGCAG 298  
Oy 1325 CCAGAACATATTTCACTGAGAACCTGACAAATCAAAAGTGCGCCAGAGGCTGTCTGTGC 1384  
Db 297 CCAGAACATATTTCACTGAGAACCTGACAAATCAAAAGTGCGCCAGAGGCTGTCTGTGC 238  
Oy 1385 TGTGCGAGCAAAAGGCTCATCCCAAAATTTGCTTACCTGATGAGGCTGTCTGTGC 1444  
Db 237 TGTGCGAGCAAAAGGCTCATCCCAAAATTTGCTTACCTGATGAGGCTGTCTGTGC 178  
Oy 1445 CTCATCCAGAGCATTTGTGCGCAAGCTTGACCCAGTCCGAGCATATGTAATCTGTG 1504  
Db 177 CTCATCCAGAGCATTTGTGCGCAAGCTTGACCCAGTCCGAGCATATGTAATCTGTG 118  
Oy 1505 GAGGCTTAAGTTGAGAACAGAGCTCTCAGCGCCAGGTGGAAGTGGCTCTCATAGCCT 1564  
Db 117 GAGGCTTAAGTTGAGAACAGAGCTCTCAGCGCCAGGTGGAAGTGGCTCTCATAGCCT 58  
Oy 1565 CCATTCGTATGAGAACAGAGCTTTCTGAAAAGGATCCAGCACACCTCTCGGTTTT 1621  
Db 57 CCATTCGTATGAGAACAGAGCTTTCTGAAAAGGATCCAGCACACCTCTCGGTTTT 1

RESULT 5  
AL535072 1065 bp mRNA linear EST 13-FEB-2001  
LOCUS AL535072 LTI\_FL013\_Fbrnl Homo sapiens cDNA clone CS0DF007YL07 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL535072  
VERSION AL535072.1 GI:12798565  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1065)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
Source  
1..1065  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF007YL07"  
/clone\_lib="LTI\_FL013\_Fbrnl"  
/dev\_stage="pooled tissue from post conception fetuses (20  
week, 24 week and 26 week)"  
/lab\_host="DH10B"  
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-stranded cDNA was digested with Not I  
and cloned into the Not I and Eco RV sites of the  
pCMVSPORT 6 vector. Library was constructed by Life  
Technologies. Contact : Feng Liang Life Technologies, a  
division of Invitrogen 9800 Medical Center Drive, Rockville  
Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 271 a 270 c 295 g 228 t 1 others

Query Match 62.2% Score 1028.8; DB 9; Length 1065;  
Best Local Similarity 99.0%; Pred. No. 2,6e-293;  
Matches 1055; Conservative 1; Mismatches 8; Indels 2; Gaps 2;  
Oy 16 TCGGAGACAGCGCGCGGTGCTGTGTGGCCATGCGCCGACTACCTGATTTAGTGGGGCA 75  
Db 1 TCGGAGACAGCGCGCGGTGCTGTGTGGCCATGCGCCGACTACCTGATTTAGTGGGGCA 60  
Oy 76 CGTCTACGTGCGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 135  
Db 61 CGTCTACGTGCGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 120  
Oy 136 TCACATACATGACTTTCATCTCTCTGCGGATGACATGACATGACATGACATGACATGAC 195  
Db 121 TCACATACATGACTTTCATCTCTCTGCGGATGACATGACATGACATGACATGACATGAC 180  
Oy 196 ACCTGACTTCTGCTGACACAGAAATACCTTTAAGACCCACTGTCTCTCTCCCA 255  
Db 181 ACCTGACTTCTGCTGACACAGAAATACCTTTAAGACCCACTGTCTCTCTCCCA 240  
Oy 256 TGGACACAGTCACAGAGGCTGGGATGGCCATAGCATGGGCGTTACAGGGGATTTGGCT 315  
Db 241 TGGACACAGTCACAGAGGCTGGGATGGCCATAGCATGGGCGTTACAGGGGATTTGGCT 300  
Oy 316 TCATCCAGCACAACCTGTACACCTGAATTCAGGCCAATGAAGTTCGAAAGTGAAAGAAAT 375  
Db 301 TCATCCAGCACAACCTGTACACCTGAATTCAGGCCAATGAAGTTCGAAAGTGAAAGAAAT 360  
Oy 376 ATGACAGGAGATTCATACAGACCCGTGTGCTGACCCCAAGATTCGCGTGGCGGATG 435  
Db 361 ATGACAGGAGATTCATACAGACCCGTGTGCTGACCCCAAGATTCGCGTGGCGGATG 420  
Oy 436 TTTTGGAGCCAAAGGCCCGCATGTTTCTGCGGATATCCCAATCACAGACAGCGCGCA 495  
Db 421 TTTTGGAGCCAAAGGCCCGCATGTTTCTGCGGATATCCCAATCACAGACAGCGCGCA 480  
Oy 496 TGGGAGCCCGCTGGTGGGATCATCTCTCCAGGACATTTGTTTTCTCAAGAGAGAGG 555  
Db 481 TGGGAGCCCGCTGGTGGGATCATCTCTCCAGGACATTTGTTTTCTCAAGAGAGAGG 540  
Oy 556 AACATGACTGTTTTCTGGAAGAGATATGACAAAGAGGGAAGACTTGTGGTACGCCCCC 615  
Db 541 AACATGACTGTTTTCTGGAAGAGATATGACAAAGAGGGAAGACTTGTGGTACGCCCTG 600  
Oy 616 GACGATTCACACTGAAGAGAGCAATGAATTTGTGACGCGCAGCAAGAGGAAAGTTGC 675  
Db 601 GACGATTCACACTGAAGAGAGCAATGAATTTGTGACGCGCA-CAAGAGAGGAAAGTTGC 659  
Oy 676 CCATTTGAATGAAGATGATGAGCTTGTGGCATCTATTTGCCGAGACAGCTGAAGAGA 735  
Db 660 CCATTTGAATGAAGATGATGAGCTTGTGGCATCTATTTGCCGAGACAGCTGAAGAGA 719  
Oy 736 ATCGGACTACCCACTAGCCTCCAAAGATCCAGAAGACAGCTGTGTGGGCGACCA 795  
Db 720 ATCGGACTACCCACTAGCCTCCAAAGATCCAGAAGACAGCTGTGTGGGCGACCA 779  
Oy 796 TTGGCACTCATGAGATGACAGTATAGCTGTGACTTGTGCGCCAGGCGTGGGTGGATG 855  
Db 780 TTGGCACTCATGAGATGACAGTATAGCTGTGACTTGTGCGCCAGGCGTGGGTGGATG 839  
Oy 856 TAGTGTTTTGGACTCTTCCAGGAAATTCATCTTCCAGATCAATATGATCAATAGTACA 915  
Db 840 TAGTGTTTTGGACTCTTCCAGGAAATTCATCTTCCAGATCAATATGATCAATAGTACA 899  
Oy 916 TCAGAACAAATACCTTAATTCGCAAGTATGAGGAGCAATGGTGCATGCTGCCAGG 975  
Db 900 TCAGAACAAATACCTTAATTCGCAAGTATGAGGAGCAATGGTGCATGCTGCCAGG 959  
Oy 976 CCAAGACCTCATTTGATGACGCTGTGATGCCCTGCGGGT-GGGCAATGGAGTGGCTCC 1034  
Db 960 CCAAGACCTCATTTGATGACGCTGTGATGCCCTGCGGGTGGGCAATGGAGTGGCTCC 1019

```
QY 1035 ATCTGCAATTATCCAGAGAGTGTGGCTGTGTGGCGGCCCAAGCAA 1080
|||||
Db 1020 ATCTGCAATTATCCAGAGAGTGTGGCTGTGTGGCGGCCCAAGCAA 1065
|||||

RESULT 6
AL539530 1049 bp mRNA linear EST 16-FEB-2001
LOCUS AL539530
DEFINITION AL539530 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF033YP03 5
prime, mRNA sequence.
ACCESSION AL539530
VERSION AL539530.1 GI:12868830
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1049)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..1049
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF033YP03"
/clone_lib="LTI_FL013_FBrn1"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
f.liang@life.techn.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 267 a 265 c 288 g 227 t 2 others
ORIGIN

Query Match 61.5%; Score 1017.6; DB 9; Length 1049;
Best Local Similarity 98.7%; Pred. No. 5.3e-290;
Matches 1034; Conservative 2; Mismatches 11; Indels 1; Gaps 1;
```

```
Db 301 ACACTGTACACCTTAATTCACGGCCAAATGAAGTTGGAAAGTGAAGAAATATGAACAG 360
QY 385 GATTTCATCACAGACCCCTGTGTCTCTACACCCCAAGATCCGCTGGGAGATGTTTTGAGG 444
|||||
Db 361 GATTTCATCACAGACCCCTGTGTCTCTACACCCCAAGATCCGCTGGGAGATGTTTTGAGG 420
|||||

QY 445 CCAAGGCCCGGATGTTTGTGGGATATCCCATATCACAGACACAGCCGGATGGGAGCC 504
|||||
Db 421 CCAAGGCCCGGATGTTTGTGGGATATCCCATATCACAGACACAGCCGGATGGGAGCC 480
|||||

QY 505 GCTGTGGGATCATCTTCCTCCAGGACATGATTTTCTCAAAGAGGAACATGTACT 564
|||||
Db 481 GCTGTGGGATCATCTTCCTCCAGGACATGATTTTCTCAAAGAGGAACATGTACT 540
|||||

QY 565 GTTCTTGAAGAGATTAATGACAAAGAGGAAGACTGGGTAGCCCCCGCAGCATCA 624
|||||
Db 541 GTTCTTGAAGAGATTAATGACAAAGAGGAAGACTGGGTAGCCCCCGCAGCATCA 600
|||||

QY 625 CACTAAGAGGAGCAATGAATATTTGACAGCGCAGCAGAGAGGAAGTTGCCATTGTGA 684
|||||
Db 601 CACTAAGAGGAGCAATGAATATTTGACAGCGCAGCAGAGAGGAAGTTGCCATTGTGA 660
|||||

QY 685 ATGAGATGATGAGCTGTGGCCATCATTTGCCGACAGACCTGAAGAAGATCGGACT 744
|||||
Db 661 ATGAGATGATGAGCTGTGGCCATCATTTGCCGACAGACCTGAAGAAGATCGGACT 720
|||||

QY 745 ACCCACTAGCCCTCAAGATGACCAAGAAACAGCTCTGTGTGGGAGCAGCATTTGCCACTC 804
|||||
Db 721 ACCCACTAGCCCTCAAGATGACCAAGAAACAGCTCTGTGTGGGAGCAGCATTTGCCACTC 780
|||||

QY 805 ATGAGATGACAAAGATATGAGCTGTGACTTCCTGCCAGGCTGTGTGATGTAGTGGTTT 864
|||||
Db 781 ATGAGATGACAAAGATATGAGCTGTGACTTCCTGCCAGGCTGTGTGATGTAGTGGTTT 840
|||||

QY 865 TGGACTCTTCCAGGAGAAATTCATCTTCCAGATCAATATGATCAATCAAGACA 924
|||||
Db 841 TGGACTCTTCCAGGAGAAATTCATCTTCCAGATCAATATGATCAATCAAGACA 900
|||||

QY 925 AATACCCTAATCTCCAAAGTATGAGGCAATGTGTCACCTCTGCCAGGCCAAGAAC 984
|||||
Db 901 AATACCCTAATCTCCAAAGTATGAGGCAATGTGTCACCTCTGCCAGGCCAAGAAC 960
|||||

QY 985 TCATGTATGACAGTGTGGATGGCCCTGGGGTGGGATGGGAAGTGGCTCATCTGATTA 1044
|||||
Db 961 TCATGTATGACAGTGTGGATGGCCCTGGGGTGGGATGGGAAGTGGCTCATCTGATTA 1019
|||||

QY 1045 TCCAGGAAGTGTGGCTGTGGCGGCC 1072
Db 1020 CCGAGGAAGTGTGGCTGTGGCGGCC 1047
|||||

RESULT 7
AL551909 1011 bp mRNA linear EST 16-FEB-2001
LOCUS AL551909
DEFINITION AL551909 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1060Y610 5
prime, mRNA sequence.
ACCESSION AL551909
VERSION AL551909.1 GI:12890302
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..1011
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D106010"
/clone_lib="LTL_NFL006_PL2"
/lisue_type="Placenta"
/notice="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
cloned, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

BASE COUNT 261 a 255 c 275 g 218 t 2 others

ORIGIN

Query Match 59.7% Score 987; DB 9; Length 1011;  
Best Local Similarity 99.0%; Pred. No. 6.1e-281;

Matches 1001; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

```
23 CACGGGGGGTCTCTGTTGGCCATGGCCGACACTGATGATTAGGCGGCACTCTCTA 82
1 CACGGGGGGTCTCTGTTGGCCATGGCCGACACTGATGATTAGGCGGCACTCTCTA 60
83 CGTGGCAGACGACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 142
61 CCGCCACAGACGACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
143 CAATGACTTTTCATTTCCCTGGGTACATGACTTCTACAGACCA GTGACCTGAC 202
121 CAATGACTTTTCATTTCCCTGGGTACATGACTTCTACAGACCA GTGACCTGAC 180
203 TTCTGCTGACCAAGAAATCATTAAAGACCCCACTGTTCTCTCCTC 262
181 TTCTGCTGACCAAGAAATCATTAAAGACCCCACTGTTCTCTCCTC 240
263 AGTCACAGAGGTGGGATGGCCATAGCAATGGCGTTACAGCGGTAT 322
241 AGTCACAGAGGTGGGATGGCCATAGCAATGGCGTTACAGCGGTAT 300
323 CCACAACCTGACACTGATTCAGAGCCCAATGAATGGGAAGTGA 382
301 CCACAACCTGACACTGATTCAGAGCCCAATGAATGGGAAGTGA 360
383 GGGATTCATCAGACACCTGTGTCTCTCAGCCCAAGATGCGGTGCG 442
361 GGGATTCATCAGACACCTGTGTCTCTCAGCCCAAGATGCGGTGCG 420
443 GGGCAGGCGCGGATGTTCTGCGGTATCCCAATCAGACACAGG 502
421 GGGCAGGCGCGGATGTTCTGCGGTATCCCAATCAGACACAGG 480
503 CCGCTTGGTGGGATCATCTCTCCAGGACATTGATTTTCTCAAGA 562
481 CCGCTTGGTGGGATCATCTCTCCAGGACATTGATTTTCTCAAGA 540
563 CTGTTTCTTGAAGAGATATGACAAAGAGGAGATTGGGTGAGG 622
541 CTGTTTCTTGAAGAGATATGACAAAGAGGAGATTGGGTGAGG 600
623 CACACTGAAGGAGCAATGAATTTCTCAGCGCAGCAAGAAGGAGAA 682
601 CACACTGAAGGAGCAATGAATTTCTCAGCGCAGCAAGAAGGAGAA 660
683 AATGAATGATGATGAGCTTGTGGCCATATGGCCGACAGACCTGAA 742
661 AATGAATGATGATGAGCTTGTGGCCATATGGCCGACAGACCTGAA 720
743 CTACCCACTAGCTTCAGAAAGATGCCAAGAACAGCTGCTGTGTGGG 802
721 CTACCCACTAGCTTCAGAAAGATGCCAAGAACAGCTGCTGTGTGGG 780
```

```
803 TCATAGATGACAGATATAGCTGACTTGTCTCCCAAGCGCTGTGTGATGTAGTGT 862
781 TCATAGATGACAGATATAGCTGACTTGTCTCCCAAGCGCTGTGTGATGTAGTGT 840
863 TTTGACTCTTCCCAAGGAAATTCATGTCAGATCAATATGATCAAGTATCAAGA 922
841 TTTGACTCTTCCCAAGGAAATTCATGTCAGATCAATATGATCAAGTATCAAGA 900
923 CAATATCCCTAATTCAGATGATGAGGCAATGTGTCTACTGTGCGCCAGCCAGAA 982
901 CAATATCCCTAATTCAGATGATGAGGCAATGTGTCTACTGTGCGCCAGCCAGAA 960
983 CCTCATTTATGATGAGTGGATGCGCTGCGGTGGCGCATGGAAGTGGCTC 1033
961 CCTCATTTATGATGAGTGGATGCGCTGCGGTGGCGCATGGAAGTGGCTC 1010
```

RESULT 8  
AL514250  
LOCUS 1028 bp mRNA linear EST 13-FEB-2001  
DEFINITION AL514250 LTL\_NFL006.PL2 Homo sapiens cDNA clone Cl0BB006Z01.5  
prime.mRNA sequence.  
ACCESSION AL514250  
VERSION AL514250.1 GI:12777744  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1028)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES  
source  
location/Qualifiers  
1..1028  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Cl0BB006Z01"  
/clone\_lib="LTL\_NFL006\_PL2"  
/lisue\_type="Placenta"  
/notice="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : filang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 265 a 260 c 278 g 224 t 1 others

Query Match 58.9% Score 973.4; DB 9; Length 1028;  
Best Local Similarity 99.2%; Pred. No. 6.6e-277;

Matches 999; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```
23 CACGGGGGGTCTCTGTTGGCCATGGCCGACACTGATGATTAGTGGGCGACGCTCTTA 82
1 CACGGGGGGTCTCTGTTGGCCATGGCCGACACTGATGATTAGTGGGCGACGCTCTTA 60
83 CGTGGCAGACGACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 142
61 CGTGGCAGACGACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
143 CAATGACTTTTCATTTCCCTGGGTACATGACTTCTACAGACCAAGTGGACCTGAC 202
121 CAATGACTTTTCATTTCCCTGGGTACATGACTTCTACAGACCAAGTGGACCTGAC 180
```

```
Qy 203 TTCTGCTGTGACCAAGAAATACCTTTAAGACCCACCTGTTTCTCTCCCATGAGAC 262
    |||||
Db 181 TTCTGCTGTGACCAAGAAATACCTTTAAGACCCACCTGTTTCTCTCCCATGAGAC 240
Qy 263 AGTCACAGAGGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCTTATCCA 322
    |||||
Db 241 AGTCACAGAGGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCTTATCCA 300
Qy 323 CCACAACTGTACACCTGTAATTCAGAGCCCAATGAACTGGGAAGTGAAGAATATGACA 382
    |||||
Db 301 CCACAACTGTACACCTGTAATTCAGAGCCCAATGAACTGGGAAGTGAAGAATATGACA 360
Qy 383 GGGATTTCATCACAGACCCCTGTGTCCTCAGCCCCAAGATCCGCTGCGGATGTTTTGA 442
    |||||
Db 361 GGGATTTCATCACAGACCCCTGTGTCCTCAGCCCCAAGATCCGCTGCGGATGTTTTGA 420
Qy 443 GGGCAAGGCGCGGATGTTTCTGCGGATTCCTCATCACAGACAGCGCGGATGGGAG 502
    |||||
Db 421 GGGCAAGGCGCGGATGTTTCTGCGGATTCCTCATCACAGACAGCGCGGATGGGAG 480
Qy 503 CCGCTTGGTGGGATCATCTCCTCCAGGACATGATTTTTCAAAGAGAGAGACATGA 562
    |||||
Db 481 CCGCTTGGTGGGATCATCTCCTCCAGGACATGATTTTTCAAAGAGAGAGACATGA 540
Qy 563 CTGTTTCTTGAAGATPAATGACAAAGAGGAGACTTTGTTAGCCCCCGCAGCAT 622
    |||||
Db 541 CTGTTTCTTGAAGATPAATGACAAAGAGGAGACTTTGTTAGCCCCCGCAGCAT 600
Qy 623 CACACTGAAGGAGCAAAATGAAATCTGCAGGCGAAGGAAGGAAGTCCCATTTG 682
    |||||
Db 601 CACACTGAAGGAGCAAAATGAAATCTGCAGGCGA-CAGGAAGGGAAGTGGCCATTTG 659
Qy 683 AAATGAAGATGATGACTTGTGGCCATCTTGGCCGACAGACCTGAAGAAGATCGGA 742
    |||||
Db 660 AAATGAAGATGATGACTTGTGGCCATCTTGGCCGACAGACCTGAAGAAGATCGGA 719
Qy 743 CTACCCACTAGCTTCCAAAGATGCCAAGAACAGCTGCTGTGTGGGCGAGCCATTGGCAG 802
    |||||
Db 720 CTACCCACTAGCTTCCAAAGATGCCAAGAACAGCTGCTGTGTGGGCGAGCCATTGGCAG 779
Qy 803 TCATGAGATGACAGTATGAGCTGGACTGTGCTGCCAGGCTGGTGGATGATGCTGT 862
    |||||
Db 780 TCATGAGATGACAGTATGAGCTGGACTGTGCTGCCAGGCTGGTGGATGATGCTGT 839
Qy 863 TTTGGACTCTTCCAGGGAATTCATCTCCAGATCAATATGATCAAGTACATAAGA 922
    |||||
Db 840 TTTGGACTCTTCCAGGGAATTCATCTCCAGATCAATATGATCAAGTACATAAGA 899
Qy 923 CAAATACCTTAATCTCCAAAGTCAATGAGCAATGCTGCTGCTGCCAGGCCAAGAA 982
    |||||
Db 900 CAAATACCTTAATCTCCAAAGTCAATGAGCAATGCTGCTGCTGCCAGGCCAAGAA 959
Qy 983 CCTCTTGTATGAGGATGAGTGGCTCGGGGTGGGCAATGGGAATG 1029
    |||||
Db 960 CCTCTTGTATGAGGATGAGTGGCTCGGGGTGGGCAATGGGAATG 1005

RESULT 9
AL560269 994 bp mRNA linear EST 16-FEB-2001
LOCUS AL560269 LTI_FL011_BCI Homo sapiens cDNA clone CS0D007YK18 5 prime
DEFINITION , mRNA sequence.
ACCESSION AL560269
VERSION AL560269.1 GI:12906570
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 994)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
```

```
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        1..994
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS0D007YK18"
        /clone_lib="LTI_FL011_BCI"
        /sex="male"
        /tissue_type="B cells from Burkitt lymphoma"
        /lab_host="DH10B"
        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
        with a NotI-oligo(dT) primer. Five prime end enriched,
        double-stranded cDNA was digested with Not I and cloned
        into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
        Library was constructed by Life Technologies. Contact :
        Feng Liang Life Technologies, a division of Invitrogen
        9800 Medical Center Drive Rockville, Maryland 20850, USA
        Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
        http://fulllength.invitrogen.com"

BASE COUNT 255 a 252 c 268 g 216 t 3 others
ORIGIN

Query Match 58.0%; Score 959.4; DB 9; Length 994;
Best Local Similarity 98.2%; Pred. No. 9,2e-273;
Matches 977; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

Qy 23 CACGCGGCGGTCTCTGTGTTGGCCATGGCCGACTACCTGATTAAGTGGGCGAGTCTCTA 82
    |||||
Db 1 CACGCGGCGGTCTCTGTGTTGGCCATGGCCGACTACCTGATTAAGTGGGCGAGTCTCTA 60
Qy 83 CGTCCAGACGACGAGACACAGCAGCAGCAGCTCTCAACTGGGAGAGGCGCTCACCTA 142
    |||||
Db 61 CGTCCAGACGAGCAGCTCAGCAGCAGCAGCTCTCAACTGGGAGAGGCGCTCACCTA 120
Qy 143 CAATGACTTTCATTTCCCTGGGTACATGACTTCACTCAGACGAGTGAAGTGAAGTGA 202
    |||||
Db 121 CAATGACTTTCATTTCCCTGGGTACATGACTTCACTCAGACGAGTGAAGTGAAGTGA 180
Qy 203 TTCTGCTGTGACCAAGAAATACCTTTAAGACCCACCTGTTTCTCTCCCATGAGAC 262
    |||||
Db 181 TTCTGCTGTGACCAAGAAATACCTTTAAGACCCACCTGTTTCTCTCCCATGAGAC 240
Qy 263 AGTCACAGAGGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCTTATCCA 322
    |||||
Db 241 AGTCACAGAGGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCTTATCCA 300
Qy 323 CCACAACTGTACACCTGTAATTCAGAGCCCAATGAACTGGGAAGTGAAGAATATGACA 382
    |||||
Db 301 CCACAACTGTACACCTGTAATTCAGAGCCCAATGAACTGGGAAGTGAAGAATATGACA 360
Qy 383 GGGATTTCATCACAGACCCCTGTGTCCTCAGCCCCAAGATCCGCTGCGGATGTTTTGA 442
    |||||
Db 361 GGGATTTCATCACAGACCCCTGTGTCCTCAGCCCCAAGATCCGCTGCGGATGTTTTGA 420
Qy 443 GGGCAAGGCGCGGATGTTTCTGCGGATTCCTCATCACAGACAGCGCGGATGGGAG 502
    |||||
Db 421 GGGCAAGGCGCGGATGTTTCTGCGGATTCCTCATCACAGACAGCGCGGATGGGAG 480
Qy 503 CCGCTTGGTGGGATCATCTCCTCCAGGACATGATTTTTCAAAGAGAGAGACATGA 562
    |||||
Db 481 CCGCTTGGTGGGATCATCTCCTCCAGGACATGATTTTTCAAAGAGAGAGACATGA 540
Qy 563 CTGTTTCTTGAAGATPAATGACAAAGAGGAGACTTTGTTAGCCCCCGCAGCAT 622
    |||||
Db 541 CTGTTTCTTGAAGATPAATGACAAAGAGGAGACTTTGTTAGCCCCCGCAGCAT 600
Qy 623 CACACTGAAGGAGCAAAATGAAATCTGCAGGCGAAGGAAGGAAGTGGCCATTTGT 682
    |||||
Db 601 CACACTGAAGGAGCAAAATGAAATCTGCAGGCG-SAAGCAAGGGAAGTGGCCATTTGT 659
```



Y	683	AAATACATGATGATGAGCTTGTGGCCATCATATTCGCCGACAGACCTGAA	AAACATTCGGCA	742
Db	660	AAATGAAAGATGATGAGCTGTGTGGCCATCATTCGCCGACAGACCTGAA	AAACATTCGGCA	719
Y	743	CTACCCACTAGCCTCCAAAGATGCGCAAGAACAGCTGCTGTGTGGGCC	GGCATTGGCAG	802
Db	720	CTACCCACTAGCCTCCAAAGATGCGCAAGAACAGCTGCTGTGTGGGCC	GGCATTGGCAG	779
Y	803	TCATGAGATGATGACAGTATATAGCTGTGACCTGTCTGCCGCCAGCTGTGT	GATGTAGTGT	862
Db	780	TCATGAGATGATGACAGTATATAGCTGTGACCTGTCTGCCGCCAGCTGTGT	GATGTAGTGT	839
Y	863	TTTGACATCTTTCGCCAGGAAATTCATCTTCAGATCAATATATCAAA	TACATCAAGA	922
Db	840	TTTGACATCTTTCGCCAGGAAATTCATCTTCAGATCAATATATCAAA	TACATCAAGA	899
Y	923	CAAAATACCTTAATTCGCAAGTCATTTGGAGGCAATGGCTCATGTGCGCC	CAGGCCAAGA	982
Db	900	CAAAATACCTTAATTCGCAAGTCATTTGGAGGCAATGTGCTGCTGCTGCC	CAGGCCAAGA	959
Y	983	CCTCATTTGATGACGAGTGTGATGATCCCTCGCGGTGG	1017	
Db	960	CTTCATTTGATGACGAGTGTGATGATCCCTCGCGGTGG	994	
RESULT 10				
AL516817/c				
LOCUS	AL516817	996 bp	mRNA	linear
DEFINITION	AL516817 LTI_NFL011.NBC1 Homo sapiens CDNA clone			EST 13-FEB-2001
ACCESSION	AL516817			CS0DA007YD24 3
VERSION	AL516817.1	GI:12780310		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
AUTHORS	1. (bases 1 to 996)			
TITLE	Li, W.B., Gruber, C., Jessup, J. and Polayes, D.			
JOURNAL	Full-length cDNA libraries and normalization			
COMMENT	Unpublished (2001)			
CONTACT	Contact: Genoscope			
GENOSCOPE	Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France				
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.				
FEATURES	Location/Qualifiers			
SOURCE	1..996			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="CS0DA007YD24"			
	/clone_lib="LTI_NFL011.NBC1"			
	/sex="male"			
	/tissue_type="neuroblastoma cells"			
	/lab_host="DH10B"			
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand CDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded CDNA was digested with Not I and Eco RV site vector. Library is not normalized, but the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Li, division of Invitrogen 9800 Medical Center Drive Rockville Maryland 20850, USA Fax : (1) 301 660 8371 Email : fliang@lifestechn.com URL : http://fulllength.invitrogen.com"			
BASE COUNT	221 a	279 c	239 g	248 t
ORIGIN				9 others

Query Match	56.0%;	Score 925.8;	DB 9;	Length 996;
Best Local Similarity	96.8%;	Pred. No. 8.4e-263;		
Matches 970; Conservative	5;	Mismatches 21;	Indels 6;	Gaps 3;

Qy	587	AAAGGGGAAACATGGTGTAGAGCCCCCGGAGCATCACTACGTAAGGAGGCAAAATGAAT	646
Db	996	AAAGGGGAAACATGGTGTAGAGCCCCCGGAGCATCACTACGTAAGGAGGCAAAATGAAT	940
Qy	647	TCTGAGCGGCACAAAGAGGAAAGTGGCCATTTGTAATGAATGATGAGCTTGTGGC	706
Db	939	TCTGGA - GGCACAAAGAGGAAAGTGGCCATTTGTAATGAGATGATGAGCTTGTGGC	881
Qy	707	CATCATTTCCCGGAGACAGACTGGAAGAAATTCGGGACTACCCACTAGAGCTCCAAAGATGC	766
Db	880	CATCATTTCCCGGAGACAGACTGGAAGAAATTCGGGACTACCCACTAGAGCTCCAAAGATGC	821
Qy	767	CAAGAACAGCTGCTGTGTGGGCGACGCCATTGGCACTATGAGGATGACAACTATAGCT	826
Db	820	CAAGAACAGCTGCTGTGTGGGCGACGCCATTGGCACTATGAGGATGACAACTATAGCT	761
Qy	827	GCAGTTCGTCCGACAGGCTGGGTGTGATGTAGTGTTTTGGACTCTTCCAGGAAATTC	886
Db	760	GCAGTTCGTCCGACAGGCTGGGTGTGATGTAGTGTTTTGGACTCTTCCAGGAAATTC	701
Qy	887	CATCTTCAGATCAATATGATCAAGTACATCAAGACAAATACCTTAATCTTCAAGTCAT	946
Db	700	CATCTTCAGATCAATATGATCAAGTACATCAAGACAAATACCTTAATCTTCAAGTCAT	641
Qy	947	TGGAGGCAATGGGTCATCTGTCGCCACGACCAAGAACCTCATGATGAGGTGTGATGC	1006
Db	640	TGGAGGCAATGGGTCATCTGTCGCCACGACCAAGAACCTCATGATGAGGTGTGATGC	582
Qy	1007	CTGTGGGTGGGCATGGGAGTGGCTCCATCTGCATTTATCCAGGAAGTGTGGCCTGTGG	1066
Db	581	CTGTGGGTGGGCATGGGAGTGGCTCCATCTGCATTTATCCAGGAAGTGTGGCCTGTGG	523
Qy	1067	GGGGCCCCAACACAGCAGAGTGTACAAAGGTATGATGATGACAGCGCCTTGGGTGTGC	1126
Db	522	GGGGCCCCAACACAGCAGAGTGTACAAAGGTATGATGATGATGACAGCGCCTTGGGTGTGC	463
Qy	1187	GGCGTCCACAGTCATGATGGGCTGTCTCTGGTCCGCACACTGAGGGCCCGTGGGATTA	1246
Db	402	GGCGTCCACAGTCATGATGGGCTGTCTCTGGTCCGCACACTGAGGGCCCGTGGGATTA	343
Qy	1247	CTTCTTTTCCGATGGGATCCGGCTAAAGAAATATTCGCGATATGGATTTCTTCGATGCAT	1306
Db	342	CTTCTTTTCCGATGGGATCCGGCTAAAGAAATATTCGCGATATGGATTTCTTCGATGCAT	283
Qy	1307	GGACAGACCTTCAGACAGCCGACACAGATATTTCACTGTAAGCTGACAAATCAAGTGGC	1366
Db	282	GGACAGACCTTCAGACAGCCGACACAGATATTTCACTGTAAGCTGACAAATCAAGTGGC	223
Qy	1367	CCAGGAGACTGTCTGGTGTGTGACAGGACAAAGGGCTCAATCCACAATTTGTCCCTTACT	1426
Db	222	CCAGGAGACTGTCTGGTGTGTGACAGGACAAAGGGCTCAATCCACAATTTGTCCCTTACT	163
Qy	1427	GATTGTGTGGCATCCAAACACTATGTCAGAGACATTGTGTCCCAAGAGCTTGACCCAAGTCCG	1486
Db	162	GATTGTGTGGCATCCAAACACTATGTCAGAGACATTGTGTCCCAAGAGCTTGACCCAAGTCCG	103
Qy	1487	AGCCATGATGATCTGTGGGAGCTTAAATTTGAGAAGAGAACGTCTTCAAGCCAGGTGGA	1546
Db	102	AGCCATGATGATCTGTGGGAGCTTAAATTTGAGAAGAGAACGTCTTCAAGCCAGGTGGA	43
Qy	1547	AGGTGGCGTCCATAGCTTCCATTCGATATGAGAAGCGGCTTTT	1588
Db	42	AGGTGGCGTCCATAGCTTCCATTCGATATGAGAAGCGGCTTTT	1

RESULT 11					
BM550826					
LOCUS	1061 bp	mRNA	linear	EST 20-FEB-2002	
DEFINITION	AGENCOURT_6585235 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471805				



/note="Organ: placenta; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 254 a 247 c 256 g 210 t 1 others  
ORIGIN

Query Match 55.7%: Score 920.8; DB 9; Length 968;  
Best Local Similarity 99.2%: Pred. No. 2.5e-261;  
Matches 925; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 23 CACGCGCGGTCTCTGTGTGGCCATGGCCGACTACCTGATTTAGTGCGGACGTCCTA 82  
DB 1 CACGCGCGGTCTCTGTGTGGCCATGGCCGACTACCTGATTTAGTGCGGACGTCCTA 60  
QY 83 CCGCCAGACGAGGAGCTACACAGCAGCAGCTTCACTGCGGAGAGCGGCTCACCTA 142  
DB 61 CCGCCAGACGAGGAGCTACACAGCAGCAGCTTCACTGCGGAGAGCGGCTCACCTA 120  
QY 143 CAATGACTTTTCATTTCTCCCTGGGATCATCGACTTCACGACCACTGACCTGAC 202  
DB 121 CAATGACTTTTCATTTCTCCCTGGGATCATCGACTTCACGACCACTGACCTGAC 180  
QY 203 TTCTGCTGACCAAGAAATCATCTTAAAGCCCACTGGTTCCTCCCTTGAGACAC 282  
DB 181 TTCTGCTGACCAAGAAATCATCTTAAAGCCCACTGGTTCCTCCCTTGAGACAC 240  
QY 263 AGTCACAGAGGCTGGGATGGCCATACCAATGGCGTTACAGCGGTATGGCTTACCA 322  
DB 241 AGTCACAGAGGCTGGGATGGCCATACCAATGGCGTTACAGCGGTATGGCTTACCA 300  
QY 323 CCACACTGTACACCTGATTCAGGCCAATGAGTTGGAAGTGAATTAATATGACA 382  
DB 301 CCACACTGTACACCTGATTCAGGCCAATGAGTTGGAAGTGAATTAATATGACA 360  
QY 383 GGGATTCATCACAGACCTGTGTGCTCTAGCCCCAAGATCGCGCTGATTTTGA 442  
DB 361 GGGATTCATCACAGACCTGTGTGCTCTAGCCCCAAGATCGCGCTGATTTTGA 420  
QY 443 GGGCAAGGCGCGGATGTTTGGGTTATCCCAATCACAGACAGCGGATGGGGAG 502  
DB 421 GGGCAAGGCGCGGATGTTTGGGTTATCCCAATCACAGACAGCGGATGGGGAG 480  
QY 503 CCGCTTGGTGGGATCATCTCTCCAGGACATGATTTTCTCAAGACAGAGACATGA 562  
DB 481 CCGCTTGGTGGGATCATCTCTCCAGGACATGATTTTCTCAAGACAGAGACATGA 540  
QY 563 CCGTTCTTGAAGATGATGATGACAAAGAGGAGGAGCTGTGTGAGCGCCCGGACAT 622  
DB 541 CCGTTCTTGAAGATGATGATGACAAAGAGGAGGAGCTGTGTGAGCGCCCGGACAT 600  
QY 623 CACACTGAAGAGGCAATGAATTCGACGCGACCAAGAGGAGAAATTTGCCATTGT 682  
DB 601 CACACTGAAGAGGCAATGAATTCGACGCGACCAAGAGGAGAAATTTGCCATTGT 660  
QY 683 AATGAGATGATGACTGTGGCCATCATTTGCCGACAGACACTTAATAAGATGGGA 742  
DB 661 AATGAGATGATGACTGTGGCCATCATTTGCCGACAGACACTTAATAAGATGGGA 720  
QY 743 CTAACCACTAGGCTCCCAAGATGCCAAGAACAGCTGCTGTGGGCGGCCATTGGAC 802  
DB 721 CTAACCACTAGGCTCCCAAGATGCCAAGAACAGCTGCTGTGGGCGGCCATTGGAC 780  
QY 803 TCATGAGATGACAAATATAGCTGAGCTTGGCTCGCCAGGCTGGTGTGATGTAGTGT 862  
DB 781 TCATGAGATGACAAATATAGCTGAGCTTGGCTCGCCAGGCTGGTGTGATGTAGTGT 840

QY 863 TTGGACTCTTCCAGGCAATTCATCTCCAGATCATATGATGATACATCAAGA 922  
DB 841 TTGGACTCTTCCAGGCAATTCATCTCCAGATCATATGATGATGATCAAGA 900  
QY 923 CAATACCCCTAATCTCCAGTCAATGAGGCA 954  
DB 901 CAATACCCCTAATCTCCAGTCAATGAGGCA 932

RESULT 13  
AL534371 953 bp mRNA linear EST 13-FEB-2001  
LOCUS AL534371 LTI\_FL013\_Fbrn1 Homo sapiens cDNA clone CS0DF004YJ13 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL534371  
VERSION AL534371.1 GI:12797864  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 953)  
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source location/Qualifiers  
1..953  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF004YJ13"  
/clone\_1ib="LTI\_FL013\_Fbrn1"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"

/note="Organ: Fetal brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 251 a 243 c 247 g 209 t 3 others  
ORIGIN

Query Match 55.5%: Score 917.4; DB 9; Length 953;  
Best Local Similarity 98.8%: Pred. No. 2.5e-260;  
Matches 932; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 34 GTCTGTGTTGGCCATGCGGACTACCTGATTAAGTGGGCGAGTCTAGTCCAGAGC 93  
DB 1 GTCTGTGTTGGCCATGCGGACTACCTGATTAAGTGGGCGAGTCTAGTCCAGAGC 60  
QY 94 ACGGACTCAGACGACGAGCTCTTCAACTGGGAGAGGCGCTCAGCTCAATGACTTTC 153  
DB 61 ACGGACTCAGACGACGAGCTCTTCAACTGGGAGAGGCGCTCAGCTCAATGACTTTC 120  
QY 154 TCATTTCCCTGGGATACATGACTTCAGACGACGAGTGAACCTGACTTGTCTGA 213  
DB 121 TCATTTCCCTGGGATACATGACTTCAGACGACGAGTGAACCTGACTTGTCTGA 180  
QY 214 CCAAGAAATTCATCTTAAGACCCCACTGTTCTCTCCATGAGACAGTACAGAGG 273  
DB 181 CCAAGAAATTCATCTTAAGACCCCACTGTTCTCTCCATGAGACAGTACAGAGG 240  
QY 274 CTGGGATGGCCATGACCAATGAGCGGCTTACAGCGGATTAAGTGGCTTACCCCAACTGTA 333

Db 241 CTGGGATGGCCATGACCAATGGCCCTTACAGGGGGTATGAGCTTCATCCACACACTGTA 300  
Qy 334 CACCTGATTTCCAGGCGCAATGAAGTTGGAAGTGAAGAAATATGAACGGGATTCATCA 393  
Db 301 CACCTGATTTCCAGGCGCAATGAAGTTGGAAGTGAAGAAATATGAACGGGATTCATCA 360  
Qy 394 CAGACCCCTGTGCTCCAGCCCCCAGAGATCGCTGGGGATGTTTTTGAAGCCCAAGCCCC 453  
Db 361 CAGACCCCTGTGCTCCAGCCCCCAGAGATCGCTGGGGATGTTTTTGAAGCCCAAGCCCC 420  
Qy 454 GGCATGTTTCTGCGGTATCCCAATACAGACACAGCCGGATGGGAGCCGCTTGTGTG 513  
Db 421 GGCATGTTTCTGCGGTATCCCAATACAGACACAGCCGGATGGGAGCCGCTTGTGTG 480  
Qy 514 GCATCATCTCTCCAGGACATGATTTTCTCAAGAGAGGAGCAACATGACTGTTCTTG 573  
Db 481 GCATCATCTCTCCAGGACATGATTTTCTCAAGAGAGGAGCAACATGACTGTTCTTG 540  
Qy 574 AAGAGTAAATGACAAAGAGGGAAGACTTGTGTAGCCCCCGCAGCATTCACACTGAAG 633  
Db 541 AAGAGTAAATGACAAAGAGGGAAGACTTGTGTAGCCCCCGCAGCATTCACACTGAAG 600  
Qy 634 AGGCAATAAATTTCTGCAAGCGGACAGCAAGAGGAAAGTTCCCATTTGTAATGAAGATG 693  
Db 601 AGGCAATAAATTTCTGCAAGCGGACAGCAAGAGGAAAGTTCCCATTTGTAATGAAGATG 659  
Qy 694 ATGAGCTTGTGGCATCATTTGCCCGAGACCTGTAAGAAATTCGGGACTACCCACTAG 753  
Db 660 ATGAGCTTGTGGCATCATTTGCCCGAGACCTGTAAGAAATTCGGGACTACCCACTAG 719  
Qy 754 CCTCCAAAGATGCCAAGAAACAGCTCTGTGTGGGCGAGCCATTGGCACTCATGAGAGATG 813  
Db 720 CCTCCAAAGATGCCAAGAAACAGCTCTGTGTGGGCGAGCCATTGGCACTCATGAGAGATG 779  
Qy 814 ACAAGTATGGCTGGGAGCTGTGCTGCCAGGCTGTGATGATGATGTTTGGACTCTT 873  
Db 780 ACAAGTATGGCTGGGAGCTGTGCTGCCAGGCTGTGATGATGATGTTTGGACTCTT 839  
Qy 874 CCCAGGAAATTCATCTTCCAGATCAATATGATCAAGTACATCAAGCAAGCAATACCTTA 933  
Db 840 CCCAGGAAATTCATCTTCCAGATCAATATGATCAAGTACATCAAGCAAGCAATACCTTA 899  
Qy 934 ATCTCCAAGTCATGGAGGCAATGTGTGCTACTGCTGCCAGGC 976  
Db 900 ATCTCCAAGTCATGGAGGCAATGTGTGCTACTGCTGCCAGGC 942

RESULT 14  
BQ278808 1020 bp mRNA linear EST 07-MAY-2002  
LOCUS BQ278808  
DEFINITION AGEMCOURT\_7049427 NIH\_MGC\_107 Homo sapiens cDNA clone IMAGE:5805460  
5', mRNA sequence.  
ACCESSION BQ278808  
VERSION BQ278808.1 GI:20489016  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1020)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2044 row: 1 column: 05

FEATURES  
source High quality sequence stop: 761.  
Location/Qualifiers  
1..1020  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5805460"  
/clone\_lib="NIH\_MGC\_107"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="organ: breast; Vector: pOT8; Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
ling Hong in the Laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC library." 1 others

BASE COUNT 252 a 240 c 291 g 236 t

ORIGIN

Query Match 55.1%; Score 911; DB 14; Length 1020;  
Best Local Similarity 96.5%; Pred No. 2.1e-258;  
Matches 963; Conservative 0; Mismatches 31; Indels 4; Gaps 3;

Qy 633 GAGGCAATGAATTTCTGCAGGCGACCAAGGGAAGTTGCCATTTGTAATGAAGAT 692  
+  
Db 22 GGGGCAATGAATTTCTGAGCGACCAAGAGGAAAGTTGCCATTTGTAATGAAGAT 81  
Qy 693 GATGACCTTGTGGCATTCATTTGCCCGGACAGACCTGTAAGAAATTCGGGACTACCCACTTA 752  
+  
Db 82 GATGACCTTGTGGCATTCATTTGCCCGGACAGACCTGTAAGAAATTCGGGACTACCCACTTA 719  
Qy 753 GCCTCCAAAGATGCCAAGAAACAGCTCTGTGTGGGCGAGCCATTGGCACTCATGAGAGAT 812  
+  
Db 142 GCCTCCAAAGATGCCAAGAAACAGCTCTGTGTGGGCGAGCCATTGGCACTCATGAGAGAT 779  
Qy 813 GACAAGTATAGGCTGACTGTCTGCCAGGCTGTGTGATGATGATGTTTGGACTCTT 872  
+  
Db 202 GACAAGTATAGGCTGACTGTCTGCCAGGCTGTGTGATGATGATGTTTGGACTCTT 839  
Qy 873 TCCCAAGGAAATTCATCTTCCAGATCAATATGATCAAGTACATCAAGCAAGCAATACCTTA 932  
+  
Db 262 TCCCAAGGAAATTCATCTTCCAGATCAATATGATCAAGTACATCAAGCAAGCAATACCTTA 899  
Qy 933 AATCTCCAAGTCATGGAGGCAATGTGTGCTACTGCTGCCAGGCATTCATGAT 976  
+  
Db 322 AATCTCCAAGTCATGGAGGCAATGTGTGCTACTGCTGCCAGGCATTCATGAT 942  
Qy 993 GCAGGTGTGATGCCCTGCGGGTGGGCTATGGAAGTGGCTCATCTGATTCAGAGAA 1052  
+  
Db 382 GCAGGTGTGATGCCCTGCGGGTGGGCTATGGAAGTGGCTCATCTGATTCAGAGAA 1019  
Qy 1053 GTGTGCGCTGTGGGGGGGGCCCAAGCAAGCAAGTGTAAAGTGTATGATGACAGG 1112  
+  
Db 442 GTGTGCGCTGTGGGGGGGGCCCAAGCAAGCAAGTGTAAAGTGTATGATGACAGG 501  
Qy 1113 CGCTTTGGTTCGCGTCATCTCTGATGAGGAATCAAAATGTGGGTATATTCAGAAA 1172  
+  
Db 502 CGCTTTGGTTCGCGTCATCTCTGATGAGGAATCAAAATGTGGGTATATTCAGAAA 561  
Qy 1173 GCGTTGGCCCTTGGGGGCTCCACAGTCATGATGGGCTCTCTGCTGCTCCACACTGAG 621  
+  
Db 562 GCGTTGGCCCTTGGGGGCTCCACAGTCATGATGGGCTCTCTGCTGCTCCACACTGAG 621  
Qy 1233 GCGCGTGTGATTTCTTTTCCGATGGGATCCGGCTTAAAGAAATATTCGCGTATGGGT 1292  
+  
Db 622 GCGCGTGTGATTTCTTTTCCGATGGGATCCGGCTTAAAGAAATATTCGCGTATGGGT 681  
Qy 1293 TCCCTGATGGCATGAGCAAGACCTCAGCAGCAGCAAGCAAGATATTTTCAGTGAAGCTGAC 1352  
+  
Db 682 TCCCTGATGGCATGAGCAAGACCTCAGCAGCAGCAAGCAAGATATTTTCAGTGAAGCTGAC 741  
Qy 1353 AAATTCAAAGTGGGCCA-GGAGTGTCTGTGCTGTGACAGCAAGAGGTCAATCCAGAA 1411

|||||  
Db 742 AAATATGAAGTGGCCAGGGAGTGTCTGTCTGTGCGAGACAAAGGCTCAATCCACAA 801  
Qy 1412 ATTTGTCCTTACCTGATTTGCTGGCATCCAAACATCATGCCAGACATTTGTTGCCAAG 1471  
Db 802 ATTTGTCCTTACCTGATTTGCTGGCATCCAAACATCATGCCAGACATTTGTTGCCAAG 861  
Qy 1472 CTTGACCCAGTCCGAGCATGATGATCTGTGGGAGCTTAAG--TTTGAAGAGAACG 1529  
Db 862 CTTGACCCAGTCCGAGCATGATGATCTGTGGGAGCTTAAGTTGAAGAGAACG 921  
Qy 1530 TCTCAGCCAGTGGAA-GGTGGCTGCATACCTCCATTCCTATGAPAAAGCGCTTTT 1588  
Db 922 TCTCAGCCAGTGGAGGAGGTGGCTGCATACCTCCATTCGATGCAAGCGGTTTTT 981  
Qy 1589 CTGAAGAAGGATCCAGACACCTCTCGGTTT 1626  
Db 982 CTTGAAGAAGGATCCAGACACCTCTCGGTTT 1019

RESULT 15  
LOCUS B0058198 1023 bp mRNA linear EST 29-MAR-2002  
DEFINITION AGENCOURT\_6924369 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5814016  
ACCESSION B0058198  
VERSION B0058198.1 GI:19817538  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1023)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Lou Staudt

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LINC2066 row: p column: 17  
High quality sequence stop: 726.  
Location/Qualifiers

FEATURES  
Source 1..1023

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5814016"  
/clone\_lib="NIH\_MGC\_99"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:  
EcoRI; CDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(g). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hui in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-CDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH-MGC  
Library."

BASE COUNT 264 a 260 c 278 g 220 t 1 others

Query Match 55.0%; Score 910; DB 14; Length 1023;  
Best Local Similarity 95.1%; Pred. No. 4.1e-258;  
Matches 961; Conservative 0; Mismatches 46; Inpds 4; Gaps 2;

Qy 12 GTCTCGAGACAGCGGCGGTCTCTGTGTGGCCATGCGCGACTACTGATTACTGGG 71  
|||||

Db 1 GTCTCGAGACAGCGGCGGTCTCTGTGTGGCCATGCGCGACTACTGATTACTGGG 60  
Qy 72 GGCACGTCTTACCTGCGACAGACGAGACTTCACAGCAGACAGCTTTCATCTCGGAGAC 131  
Db 61 GGCACGTCTTACCTGCGACAGACGAGACTTCACAGCAGACAGCTTTCATCTCGGAGAC 120  
Qy 132 GGCCTACCTTACATGATTTTCATTTTCCCTGGGTACATGACTTTCATGACACAG 191  
Db 121 GGCCTACCTTACATGATTTTCATTTTCCCTGGGTACATGACTTTCATGACACAG 180  
Qy 192 GTGACCTGACTTCTCTCTGACCAAGAAATCATTAAAGCCCACTGTTCTCTCT 251  
Db 181 GTGACCTGACTTCTCTCTGACCAAGAAATCATTAAAGCCCACTGTTCTCTCT 240  
Qy 252 CCCATGACACATTCACAGAGGTGGGATGGCATATGCAATGGCGCTTACAGCGGATT 311  
Db 241 CCCATGACACATTCACAGAGGTGGGATGGCATATGCAATGGCGCTTACAGCGGATT 300  
Qy 312 GGCCTTATCCACACACACTGTACACCTGAAATTCAGAGCCAAATGAAAGTGAAG 371  
Db 301 GGCCTTATCCACACACACTGTACACCTGAAATTCAGAGCCAAATGAAAGTGAAG 360  
Qy 372 AAATATGAACAGGATTTTCATTCACAGACCTGTGTCTCTACAGCCCAAGATCGCTGGG 431  
Db 361 AAATATGAACAGGATTTTCATTCACAGACCTGTGTCTCTACAGCCCAAGATCGCTGGG 420  
Qy 432 GATGTTTTGAGCGCCAAAGCCCGGATGTTCTGCGGTATCCCATTCACAGACAGCG 491  
Db 421 GATGTTTTGAGCGCCAAAGCCCGGATGTTCTGCGGTATCCCATTCACAGACAGCG 480  
Qy 492 CGGATGGGAGACCGCTTGGTGGCATATCTCTCCAGGACATTTGTTTCAAGAG 551  
Db 481 CGGATGGGAGACCGCTTGGTGGCATATCTCTCCAGGACATTTGTTTCAAGAG 540  
Qy 552 GAGAACATGACTGTTTCTTGAAGATATATGACAAAGAGGAAACTTGGGTGAGGC 611  
Db 541 GAGAACATGACTGTTTCTTGAAGATATATGACAAAGAGGAAACTTGGGTGAGGC 600  
Qy 612 CCCCAGCATCAGCTGATGAAGAGGCAATGAATTTCTCAGCGCAGCAAGAGGAGAG 671  
Db 601 CCCCAGCATCAGCTGATGAAGAGGCAATGAATTTCTCAGCGCAGCAAGAGGAGAG 660  
Qy 672 TTGCCATTTGAATGAATGATGATGAGCTTGGCCATATTTGGCGGACAGCTGAAG 731  
Db 661 TTGCCATTTGAATGAATGATGATGAGCTTGGCCATATTTGGCGGACAGCTGAAG 720  
Qy 732 AAGATCGGACTACCCACTAGCTGCTCAAGAGATGCCAAGAACAGCTGTGTGGGCA 791  
Db 721 AAGATCGGACTACCCACTAGCTGCTCAAGAGATGCCAAGAACAGCTGTGTGGGCA 780  
Qy 792 GCCATTGGCAGCTAGGATGACAGATATGAGCTGAGCTGCTGCGCCAGAGCTGGTGG 851  
Db 781 GCCATTGGCAGCTAGGATGACAGATATGAGCTGAGCTGCTGCGCCAGAGCTGGTGG 840  
Qy 852 G---ATGATGCTGTTTGGACTCTTCCAGGGAATTCATCTTCCAGATCATATGATG 908  
Db 841 GATGCTGATGTTTGGACTCTTCCAGGGAATTCATCTTCCAGATCATATGATG 900  
Qy 909 AAGTACATCAAGACAAATATCCCTAATCTTCAAGTCAATGAGGAGGCAATTTGGGACTGT 968  
Db 901 AAGTACATCAAGACAAATATCTAATCTTCAAGTCAATGAGGAGGCAATTTGGGACTGT 960  
Qy 969 GCCCAGGCAAGAACCTCAATG-ATGCAAGGTGTGATGCTCGCGGTGGG 1018  
Db 961 GCCCAGGCAAGAACCTCTTGAATCCAGTGTGAAGAGCTTGGCGGGGG 1011

Search completed: February 12, 2003, 19:26:36  
Job time : 2260 secs

THIS PAGE BLANK (USPTO)

Page 1

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

Db 1 MADYLLISGTSYVDDGTLTAQQLFNCGDGLTYNDFLLIPGYIDFTADQVDTLSALTJKKIT 60
QY 61 LKTPLVSSPMDTVTEAGMAIMALTGGIGFIHNCCTPEFOANEVKKYKKGQFTDPV 120
    |||||||
Db 61 LKTPLVSSPMDTVTEAGMAIMALTGGIGFIHNCCTPEFOANEVKKYKKGQFTDPV 120
QY 121 LSPKDRVDVEAKARHGFCGIPITDTGRMGSRVLVGIISSRDIDFLKEEHDCELEEIMT 180
    |||||||
Db 121 LSPKDRVDVEAKARHGFCGIPITDTGRMGSRVLVGIISSRDIDFLKEEHDCELEEIMT 180
QY 181 KREDLVVAPRSITLKEANETLQRSKKGKLPVNEDELVAITARTDKKNDYPLASKDA 240
    |||||||
Db 181 KREDLVVAPRSITLKEANETLQRSKKGKLPVNEDELVAITARTDKKNDYPLASKDA 240
QY 241 KKOLCGAIAITHEDDKRYRDLDLAQAQVYVLLDSOGNSIFQIMIKYIKDKYPNLOVI 300
    |||||||
Db 241 KKOLCGAIAITHEDDKRYRDLDLAQAQVYVLLDSOGNSIFQIMIKYIKDKYPNLOVI 300
QY 301 GGNVYTAQAQKNLIDAGVDALRVGMSGSCICITQEVLAGCRPQATAYKYVEYARREGVP 360
    |||||||
Db 301 GGNVYTAQAQKNLIDAGVDALRVGMSGSCICITQEVLAGCRPQATAYKYVEYARREGVP 360
QY 361 VIADGIONVGHIAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
    |||||||
Db 361 VIADGIONVGHIAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
QY 421 DKHLSSQNRYSSEADKIKVAQVSGAVODKSGIHKEVYPIYLAGIQHSQCODIGAKSLTOVR 480
    |||||||
Db 421 DKHLSSQNRYSSEADKIKVAQVSGAVODKSGIHKEVYPIYLAGIQHSQCODIGAKSLTOVR 480
QY 481 AMMYSGLKFEKRTSSAQVEGCVHSLHSYERLRF 514
    |||||||
Db 481 AMMYSGLKFEKRTSSAQVEGCVHSLHSYERLRF 514

RESULT 2
US-08-925-230-8
; Sequence 8, Application US/08925230
; Patent No. 6147194
; GENERAL INFORMATION:
; APPLICANT: Collart, Frank
; APPLICANT: Huberman, Eliezer
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO IMPDH
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,230
; FILING DATE: September 8, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,665,583
; FILING DATE: 12-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilsoo, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ. ID NO.: 8:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 514 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
US-08-925-230-8

Query Match
Best Local Similarity 97.7%; Score 2559; DB 4; Length 514;
Matches 504; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MADYLLISGTSYVDDGTLTAQQLFNCGDGLTYNDFLLIPGYIDFTADQVDTLSALTJKKIT 60
Db 1 MADYLLISGTSYVDDGTLTAQQLFNCGDGLTYNDFLLIPGYIDFTADQVDTLSALTJKKIT 60
QY 61 LKTPLVSSPMDTVTEAGMAIMALTGGIGFIHNCCTPEFOANEVKKYKKGQFTDPV 120
    |||||||
Db 61 LKTPLVSSPMDTVTEAGMAIMALTGGIGFIHNCCTPEFOANEVKKYKKGQFTDPV 120
QY 121 LSPKDRVDVEAKARHGFCGIPITDTGRMGSRVLVGIISSRDIDFLKEEHDCELEEIMT 180
    |||||||
Db 121 LSPKDRVDVEAKARHGFCGIPITDTGRMGSRVLVGIISSRDIDFLKEEHDCELEEIMT 180
QY 181 KREDLVVAPRSITLKEANETLQRSKKGKLPVNEDELVAITARTDKKNDYPLASKDA 240
    |||||||
Db 181 KREDLVVAPRSITLKEANETLQRSKKGKLPVNEDELVAITARTDKKNDYPLASKDA 240
QY 241 KKOLCGAIAITHEDDKRYRDLDLAQAQVYVLLDSOGNSIFQIMIKYIKDKYPNLOVI 300
    |||||||
Db 241 KKOLCGAIAITHEDDKRYRDLDLAQAQVYVLLDSOGNSIFQIMIKYIKDKYPNLOVI 300
QY 301 GGNVYTAQAQKNLIDAGVDALRVGMSGSCICITQEVLAGCRPQATAYKYVEYARREGVP 360
    |||||||
Db 301 GGNVYTAQAQKNLIDAGVDALRVGMSGSCICITQEVLAGCRPQATAYKYVEYARREGVP 360
QY 361 VIADGIONVGHIAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
    |||||||
Db 361 VIADGIONVGHIAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
QY 421 DKHLSSQNRYSSEADKIKVAQVSGAVODKSGIHKEVYPIYLAGIQHSQCODIGAKSLTOVR 480
    |||||||
Db 421 DKHLSSQNRYSSEADKIKVAQVSGAVODKSGIHKEVYPIYLAGIQHSQCODIGAKSLTOVR 480
QY 481 AMMYSGLKFEKRTSSAQVEGCVHSLHSYERLRF 514
    |||||||
Db 481 AMMYSGLKFEKRTSSAQVEGCVHSLHSYERLRF 514

RESULT 3
US-09-212-247C-9
; Sequence 9, Application US/09212247C
; Patent No. 6391603
; GENERAL INFORMATION:
; APPLICANT: POMPEJUS, Markus; SUELBARGER, Harald; JOEFFKEN, Hans
; APPLICANT: WOLFGANG, DOVAL, Jose Luis Revuelta; JIMENEZ, Alberto;
; TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii
; and GARCIA, Maria Angeles Santos
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,247C

```





Matches 84; Conservative 11; Mismatches 40; Indels 2; Gaps 1;

```

QY 11 SYVPDGLTMOOLFEN--CGDGLYNDPLIPGYIDPADVDLTSAITKTIITKTPYVS 68
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 15 TVAEKGLSVQELMDSKTGGTLTYNDPLVPGKIDFPSSSVLSSRTTKTITINAPVSS 74
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 69 PMDTVEAGMAIMALTGGIGTFIHNCITPEFOANEVKKVKKYEOGFTLDPVLSPKDRVR 128
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 75 PMDTVEADMAIMHALLGGIGTFIHNCITAEQAEVRRVKKYENGFINAPVVGPATVA 134
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 129 DVEAKRHGFCIPIT 145
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 135 DVRRMKNFEGFAGFPVT 151
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

RESULT 6

US-09-134-001C-5204  
Sequence 5204, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 5204

LENGTH: 333

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5204

Query Match

Best Local Similarity 12.8%; Score 336; DB 4; Length 333;

Matches 101; Conservative 53; Mismatches 130; Indels 40; Gaps 8;

```

QY 183 EDLVAPRSTLTKEANEILORSKKG---KLPVNE-----DDEVAIIARTDL----- 227
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 15 EDIOLIPNKCIVSRSCNTSVKGRTEFLPVYPMNQTVMNEELAQWFAENDVRYIMH 74
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 228 ---KKRDVPLASDAKAKKOLCGAIGTHEDDKYRLDILAQAQV--DVVYLLDSOGSIF 282
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 75 RFDEENR-IPFIKKMHAGLFASISVGKENEENFTFKLASSSLIPRYITIDIAHGSNS 133
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 283 QIMIKYIKDKYPNLQVIGNVVTAQAQKNLIDAGVDALRVGMSGSICI--IQEVLACG 340
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 134 VINMIKHKKHLPNSFVIAGNVGPREGVELFNAGADATVIGIGPVCTIKIKTGCGTG 193
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 341 RPPATVAVYKYEVARRRGVNVIADGCIQNVGHIAKALALGASTVMGSLAATTEAGEX 400
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 194 GWSLALNLCNKAARK---PIADGGIKRTHTGDAKLSIRFGATVVMGSLFAAHEESGET 250
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 401 FFSDDGIRLKKYRGMGLDAMDKHLSSQNRYSFSDAKIKVAQVSGAVQDGSIHKKFPYL 460
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 251 VELDGKKYKEFGSASEYOKGEH-----KNVEGKKMFVEHKGSLKDTLTEM 296
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 461 IAGIOHSCODIGAKSLTOVRAMY 484
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 297 EDDLOSSISYAGGKDLKSLFTVDY 320
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

RESULT 7

US-08-774-169-1

Sequence 1, Application US/08774169

Patent No. 5756332

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: A NOVEL GUANOSINE MONOPHOSPHATE REDUCTASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/774,169

FILING DATE: Herewith

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0182 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 366 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Consensus

CLONE: Consensus

US-08-774-169-1

Query Match

Best Local Similarity 12.7%; Score 331.5; DB 1; Length 366;

Matches 95; Conservative 73; Mismatches 138; Indels 47; Gaps 10;

```

QY 183 EDLVAPRSTLTKEANEI-----LORSKK--GKLPV--NEDD-----ELVAIIARTDL 227
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 31 KVDLLRPKRSTLTKEANEI-----LORSKK--GKLPV--NEDD-----ELVAIIARTDL 227
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 228 ---KKNDYPLA-----SKDAKKOLCGAIGTHEDDKYRLDILAQA--GVDVYVLD 275
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 91 FTAVHRHYSLVQWQERFAGONPDLCEHL--AASSGTSSDFEQLLEOTIEAIPQVKTICLDV 148
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 276 SOGNSLFOINMIIKDYPNLQVIGNVVTAQAQKNLIDAGVDALRVGMSGSICIIE 335
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 149 ANYSCHFEVEFVADVKKRPPQHTIMAGNVVTEGMEVELLISGADITIKVIGIPESVCTTKR 208
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 336 VLAGGRPQATAVYKYEVARRRGVNVIADGCIQNVGHIAKALALGASTVMGSLAATTE 395
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 209 KTVGVYFPLSAVVECDAAHGLKHLISDGGCCSPQDVAKAFAGADPFVWGLGMLGHSE 268
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 396 APGEYFSDGIRLKKYRGMGLDAMDKHLSSQNRYSFSDAKIKVAQVSGAVQDGSIHK 455
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 269 SGGELLERDQKKYKLYFGMSSEMAKKYAGVAYE-----RASGKTVEYVFKGDVEH 321
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 456 FVYVLIAGIOHSCODIGAKSLTOVRAMMTSGELKFEKRTSSAOVEGVSLHS 508
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 322 TIRDLIGIRSTCTYVGAAKLKL-----SRRTTFIRVTOQVNPPIFS 363
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

RESULT 8

US-08-774-169-4

Sequence 4, Application US/08774169

Patent No. 5756332

GENERAL INFORMATION:



Db 337 E0ENRIF 343

RESULT 10  
US-08-441-139-2  
; Sequence 2, Application US/08441139  
; Patent No. 5773245  
; GENERAL INFORMATION:  
; APPLICANT: Wiltrop, Dr. Karl D.  
; APPLICANT: Robinson, Anne S.  
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF  
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,139  
; FILING DATE: 15-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/089,997  
; FILING DATE: 06-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8646  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ. ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 682 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-441-139-2

Query Match 4.5%; Score 117.5; DB 1; Length 682;  
Best Local Similarity 20.8%; Pred. No. 0.0082;  
Matches 83; Conservative 58; Mismatches 129; Indels 129; Gaps 20;

QY 127 VRVFEA-KAHGFCGPIPDTRGMSRLVGLISSRIDPLKEE-----EHGCF-- 174  
Db 283 VROLIAFKKHH--GIVSDNNKALAKL-----KRAEKAKRALSSQMSRIEIDSEVD 334  
QY 175 ---LEEIMT--KREDLVVAPRSITLKEANEILORSKKGKLPINVEDDELVAIIARTDLRK 229  
Db 335 GIDLSEILPFAKFEELNLDFKTKLPVEKVLDDSGLEKKDY---DDIYLVGSGSTRIRK 330  
QY 230 NRDIYPLASKDAKKQLLGAAGTGHEDDKYRLDLAAGV-----DVVVLDSSQGSNT 281  
Db 391 VQQLLESYFDGKK---ASKGINDEAVVGAAYGAGVLSCGEGVEDIYLLD----- 438  
QY 282 FOINMKIKYTKDKPNLOVIGG-----NVVTAQAQAKNIDAGVALRKMGSGSICIT 333  
Db 439 --VNALTL-----GLETTGGVMTPLIKRNTAIPTKKSQIFSTAVD----- 476  
QY 334 QEVLACGRPOATAVYKYEYARF-----GVPIVADGGIONVGHITAKALALG 380  
Db 477 -----NPTYMIKYIEBERAMSKDNILGKFEELTGIPAPRG---VPQIEVTFALD 524  
QY 381 ASTVMGSLLAATTAPGEYFSDGIRLKKYRGMSLDAMDKHLSSQNRYSADKIKYA 440

Db 525 ANGLIK---VSATDKGTGK---SESITITNDKQRLQOEIDRWVEBAEKFASDASIK-- 576  
QY 441 QGVSAVQDKGSIHKFVPIYLIAGIOHSCO-----DIGAK 474  
Db 577 ---AKVESRKNLENYA-----HSLKNQVNDLGEK 603

RESULT 11  
US-09-134-001C-4675  
; Sequence 4675, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4675  
; LENGTH: 691  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
; US-09-134-001C-4675

Query Match 4.2%; Score 109.5; DB 4; Length 691;  
Best Local Similarity 20.0%; Pred. No. 0.053;  
Matches 90; Conservative 64; Mismatches 137; Indels 159; Gaps 21;

QY 92 HHNCTPEQANEVRKRYKKYEGGFTDPVVLSPKDRDVPFEAKRRGFCGIPITDGRMG 151  
Db 299 HLOSTKNFVREELKTRRFLLNDLITQNYALAPYQDIR-----FLIYP-QDLPKHE 347  
QY 152 SRLVGISSRIDDFLKEEEDHCFLEETMTKREDLVVAPRSITLKEANEI-----LORSKK 206  
Db 348 KK-----NSSQIDYIPDFVLEQLEFHHNDLHKDLIVV-WIAKGTGLRISDVLTQNNCL 401  
QY 207 GKLPINVEDDELVAIIARTDLKKNRDYPLASKDAKKQLLGAAGTGHEDDKYRLDLAQA 266  
Db 402 AK---VNGKYSITLDAKTFVKGR-IPIDNKLA-----DIIA-- 435  
QY 267 GADVVLVLDSSQGSNIFQIMKIKYIKDKYR-----LOVIGGVNTAAQ 309  
Db 436 ---VLIADSKS-----KSTKDNPNPNYIFAAYKGRKGMPTQHMVRAHLNHLK 482  
QY 310 AKNLIDAGVD-----ALRVGMGSGSICITIOEVLACGROATAVY-KVEYAR 355  
Db 483 TKNIIDQGEIFHFKTHORRHRYAAVAKLNGGADILTIQELHLASSPEKMLRAKLDDRK 542  
QY 356 REGVPII-----ADGCIQNVGHIA---KAL-----ALGASTVMGSLAATT- 394  
Db 543 KRAFESVIDGAFSFDVQDKIKNIQHSSELSKALMSQEHKLNMDPYGCHARLSG 602  
QY 395 -----EAPG-----EYFSDGIRLKY-----RGMSLDAMK 422  
Db 603 DCPYMEAPPLCNSGKPKCKDLAIGFSD-LDVEKKYELHITKTVKSIELAKNNNRQDMVER 661  
QY 423 HUSSONRY-----FSEADKIV 439  
Db 662 HINILNKYEELIGNTKDGNIIFGRNRIKY 691

RESULT 12  
US-09-134-001C-3866  
; Sequence 3866, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
;; FILE REFERENCE: GTC-007  
;; CURRENT APPLICATION NUMBER: US/09/134,001C  
;; CURRENT FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/064,964  
;; PRIOR FILING DATE: 1997-11-08  
;; PRIOR APPLICATION NUMBER: US 60/055,779  
;; PRIOR FILING DATE: 1997-08-14  
;; NUMBER OF SEQ ID NOS: 5674  
;; SEQ ID NO: 3866  
;; LENGTH: 785  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3866

Query Match 4.1% Score 106.5; DB 4; Length 785;  
Best Local Similarity 22.6%; Pred. No. 0.13;  
Matches 97; Conservative 51; Mismatches 171; Indels 111; Gaps 19;

QY 115 ITDPVVLSPKRDYRVFEAKARHFGCIPITDTGRMGSRV-----GTSRRDIDLEKEE 169  
DB 189 LSDAIVTVNDRVIVPVKAEYRQDFMGI-VHDQASGOTLYIEPNSVVMNNOISRLND 247  
QY 170 EHDCELEIMTKREDLVAPRSITLKEANBIL-----QRSGKLPVNEDE 217  
DB 248 E-AVERERILTELITGLVSA-ESDALLVAESVMQIDFLAKARYARTITGKPTFRED-- 303  
QY 218 LVAIIARTLKKRNDYPLASK-----DAKKOLCGAIGTHEDDKYRIDL-- 262  
DB 304 -----RTIYLPNAFHPHLDKDTIVANTIEFIDVETVITIGTNGKVTLKTLGLIIV 357  
QY 263 LAQAGDVVVLDSOGNSIFQINMIKIKDKYPNLQVIGCNVYT-AAOKNLIADAVDL 321  
DB 358 MASGLITFTLDSQ-LSTFE-NVYCDIGDE-----OSTQSSSTFSSHKKNVEIILQDA- 410  
QY 322 RVGMSGISCIITIEVLACGRPO--ATAVYKAYEYARFGVPVLA-----DQ 365  
DB 411 ----DQNSLILFDELGAGTDPSEGALAMSLDYVRRLGSLVATTHYELKAYSYNRGG 466  
QY 366 -----GIQNGHIAKALALGASIVMMSLAAATTE-----AP 397  
DB 467 VMANAVEFDVTLSPYKLLMGVPGRSNAPDISKRGLSLNIIINAKATIGTDEQINAM 526  
QY 398 GEYFFSDGIRLKKYR-----GMSGSLDAMDKHLSSONRY--FSEDKIKVAGVSG 445  
DB 527 IESLEHNSKRVQQRIELDRLVREAQETHDALSKQYQYQVNTKSLMEIAKEKANGRVKS 586  
QY 446 AVQDKSGIHK 455  
DB 587 ATKADEILK 596

RESULT 13  
US-09-377-557-4  
; Sequence 4, Application US/09377557  
; Patent No. 6287055  
; GENERAL INFORMATION:  
; APPLICANT: FALCO, S. CARL  
; APPLICANT: FAMODU, LAYO O.  
; APPLICANT: OROZCO, EMIL M. JR.  
; TITLE OF INVENTION: Amino Acid Decarboxylases  
; FILE REFERENCE: BB-1237  
; CURRENT APPLICATION NUMBER: US/09/377,557  
; CURRENT FILING DATE: 1999-08-19  
; EARLIER APPLICATION NUMBER: 60/099,493  
; EARLIER FILING DATE: September 8, 1998  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO: 4  
; LENGTH: 484  
; TYPE: PRT  
; ORGANISM: Zea mays

US-09-377-557-4

Query Match 3.9% Score 103; DB 4; Length 484;  
Best Local Similarity 18.1%; Pred. No. 0.13;  
Matches 94; Conservative 77; Mismatches 159; Indels 188; Gaps 25;

QY 4 YLISGTSYVPDGLTAQOLFNCGDGLTYNDFLIPGYIDFTADQVLTALTKITFKT 63  
DB 79 FLVNGTTC-----GLOASVMATCSPG---DYLIYPRNCHLSVIALVSGVPKRYI-- 127  
QY 64 PLVSPMDTVTEGMAIIMALTCGIFHNNCTPREQANVEYKRYKYGCGITLDPVYSLP 123  
DB 128 PEYNSGWD-----IAGGT-----TP-LQDDEAVKELEDGKRVGAVLVTSIP 167  
QY 124 KDRVDVFEAKARHFGCIPITDTGRMGSRVIGIIS---SRDIDLEKEE-----HDC 173  
DB 168 -----YHGYC-----SNVOGIVSVCHPRGIPVIVDEAHGAFHFHDS 205  
QY 174 FLEINTKREDLVAPR-----SIT-----LKEANBILQRSGKKGLPIV 212  
DB 206 LPSTALEQADALAVOSTHAKVLCSTLQSSMLHMSGDLVDVDKYSQCLQLQSSPSYLLS 265  
QY 213 NEDELVAIIARTDLKKNR---DYPLASKDAKKOLCGAIGHEDDKYRLDLAQ-AGV 268  
DB 266 SLD-----AARDQLSQNNIFDEPLAIASETK-----DLIARIPGI 301  
QY 269 DVVVLDSOGNSIFQINMIKIKDKYPNLQVIGCNVYTAQAOKNLIDAGVDALRVGMSG 328  
DB 302 SYVDLPC-----FASDFP-----AIDPLRITLSAS 326  
QY 329 SICI-----IQEVLACGRPOATAVYKAYEYARFGVPVIADGGIQNGHIAKALALGASIV 384  
DB 327 DLQLSYEADDL-----YE-----GHQIVE--LVGTRATVPANLCTRPQ 366  
QY 385 MMSGLIAATTEAPGEYFFSDGIR-LKKYRMSGSLDAMDKHLSSONRYSEADKIKYAO-- 441  
DB 367 DAEKLVQSAKHLSSEKHFANSLKPYKENVHGPLENISVLSPREAFPEKBRVKIJDLSL 426  
QY 442 -----GVSGAVQDKSGIHKFPVPLIAGIOH 466  
DB 427 GEICGELICPPPGIPVLIIPGEVYTHDSLSTYLSMS-VRH 463

RESULT 14

US-08-441-139-7  
; Sequence 7, Application US/08441139  
; Patent No. 5773245

;; GENERAL INFORMATION:  
;; APPLICANT: WILTRUP, DR. KARL D.  
;; APPLICANT: ROBINSON, ANNE S.  
;; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF  
;; NUMBER OF SEQUENCES: 20  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
;; STREET: 400 Garden City Plaza  
;; CITY: Garden City  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 11530  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/441,139  
;; FILING DATE: 15-MAY-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/089,997  
;; FILING DATE: 06-JUL-1993  
;; ATTORNEY/AGENT INFORMATION:

NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8646  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SAMS UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-441-139-7

Query Match 3.9%; Score 102.5; DB 1; Length 663;  
Best Local Similarity 20.2%; Pred. No. 0.24; Indels 117; Gaps 23;  
Matches 107; Conservative 78; Mismatches 227;

QY 37 ILPGYIDFTADQ-----VDLTSALTFRKITLKT---PLVSSPMDTDT 74  
DB 68 ITPSYAFTFEDERLVEAKNAQAPSPENTITDIKRLGRKDEKTMADISFPHIYN 127  
QY 75 EAGMALAMALTGGIGFIHNCIP-EFQANEVRKVKYEGGFTDPVYLSPKDRVDVFEA 133  
DB 128 DKRPLVEVNVG---KKKFTPEISAMILSKMOTAEAYLGKPYT----- 171  
QY 134 KRRHGCCGIP-ITDGRNGSRVGISSRDIDFLKEEHDCLEFIMTKRE----- 183  
DB 172 ---HSVYVTPAYFNDQROATKAGTACGLNVIATVNEPTAAIAVGLDKTDEKHIVY 228  
QY 184 DLVVAPRSTLTKEAN---EILQSKKCKPLVINEDEDELVAIARTDKKNDYPLASKD 239  
DB 229 DLGGGFVDVSLISDINGVEVLATSGDTHLGDEDFDNRVINYTLARIYNNK-----NND 282  
QY 240 AKKQLLCAIG--THEDKRYDLIAQAGVDVYVLDSSQNSIFO-IMIKYIKDKYIN 296  
DB 283 VTKDL---KAMGKLRKEVEKANGTILSSQKSVRIETESFPNGODFSETLSRAKFEETKHGS 339  
QY 297 LQVIGSNVNTAAQAKNLIDAGVDALRVGSGSICIIQEVLC--GRQATAVYK----V 350  
DB 340 LQDEDFEPVOVLKDSNLKSEIDIDVLVGSGSTRIPKVOELSEFQKAKSKGINPEAVA 399  
QY 351 YEVARFEGVPIADGSIQNVGHI-AKALALGASTV--MMGSLLAATTEAP--GEYFFSDG 405  
DB 400 YCAAVQAGV-LSGEGSDNIIVLLDVIPLTIGETTGCVMTKLIGRNTPIPTKRSQFSTA 458  
QY 406 IRLKTRYGKSLDAMDKHLSSQNRITFEADKIKVAGVSGAVODKSIHKF---VPYLI 461  
DB 459 V-----DNQNTVL---IQVEGERTLTAKDNNLLCKFRLRGIPAP 495  
QY 462 AGIQ--HSCQDIGAKSLQVRAAMYSGELKPEK-----RTSSQVQ 500  
DB 496 RQVPLEITFEVDANGVLTVSAVDSKGRPEKLVIAKDKGRLESDIE 544

RESULT 15  
US-09-066-047-8  
Sequence 8, Application US/09066047A  
Patent No. 6306394  
GENERAL INFORMATION:  
APPLICANT: MURPHY, Cheryl  
STOREY, James  
BELTZ, Gerald A.  
COUGHLIN, Richard T.  
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND METHODS OF  
USE OF GRANULOCYTIC ERHLICHTIA  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: HALE AND DORR LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts

COUNTRY: United States  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,047A  
FILING DATE: 24-Apr-1998  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/044,869  
FILING DATE: 25-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Superko, Colleen  
REGISTRATION NUMBER: 39,850  
REFERENCE/DOCKET NUMBER: 106,941,156  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 526-6000  
TELEFAX: (617) 526-5000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-066-047-8

Query Match 3.9%; Score 102; DB 4; Length 420;  
Best Local Similarity 21.3%; Pred. No. 0.13;  
Matches 85; Conservative 65; Mismatches 103; Indels 146; Gaps 23;

QY 158 ISSRDIPLKEE-----EHDCELEIMTK--REDLV--VAPRSITLKEANEI 200  
DB 22 ISSKFDLIEDITQALDADVDNLGVDFEIEIVKSKIYGVGVKGLPEQVIRIEEC 81  
QY 201 L-----QSKP--RGKLP-----YNEDEDELVAIARTDLKKNDYPL----- 235  
DB 82 LIEVLGNEKSLDLDGKIPAVIMVGLQGVGKTTVTVAALR--LTKDSKNPLVASLDVY 139  
QY 236 --ASDKAKQLLCAIGTHE--DDKRYRLDLAQA-----GVDVYVLDSSQNSIFO- 283  
DB 140 RPAARQQLKVLADGVGIDSLPIVEQOKPLDIKKRAMREARLKHGVDVLDLTGRHLIND 199  
QY 284 -IMIKYIKDKYINLQVIGSNVNTAAQAKNLIDAGVDALRVGSGSICITOE----- 335  
DB 200 MIDEKLCVKEVSPAEIV-----LV---VDSL---MGODAVYVVKFNEDELGI 241  
QY 336 -----VLACGRQATAVYKVEYARFQVY--IADG-----GIQNVGH 372  
DB 242 TGTITRADGDRGGALISMKLVA---GCPILKFMSTGKPEDLDDFYDPRIARRNLMGD 298  
QY 373 IAKALALGASTVMMGSLLAATTEAPGEYFFSDGIRLKYRGKSLDAMDKHLSSQNRIFS 432  
DB 299 VA-----SLVKAVEAVGKDTINE-LQAKKAKKGFDDLDLVYQLKALNK--- 341  
QY 433 EADKIKVAGVSGAVODKSIHKFVPLY-----IAGI 464  
DB 342 -----MGIT---ANIMKFTIPAFGNDIKRRVAGI 366

Search completed: February 13, 2003, 06:34:39  
Job time : 43 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2003, 06:32:42 Search time: 04 Seconds  
(without alignment)  
43.198 Million cell updates/sec

Title: US-09-846-637C-4

Sequence: 1 MADYLLISGTSYVPDDGLTA.....SSAQGVGVHLSYKRLRF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database: Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by charge to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2605	99.5	514	10	US-09-853-918-63
2	2595	99.1	514	10	US-09-853-918-49
3	2234	85.3	514	10	US-09-853-918-62
4	2222	84.8	514	10	US-09-853-918-48
5	2202	84.1	514	10	US-09-853-918-64
6	2201	84.0	514	10	US-09-853-918-65
7	1862	71.1	384	10	US-09-853-918-20
8	1861.5	71.1	385	10	US-09-853-918-34
9	1860.5	71.0	385	10	US-09-853-918-32
10	1860	71.0	384	10	US-09-853-918-29
11	1859.5	71.0	385	10	US-09-853-918-26
12	1858	70.9	384	10	US-09-853-918-37
13	1857.5	70.9	385	10	US-09-853-918-34
14	1857.5	70.9	385	10	US-09-853-918-38
15	1857	70.9	384	10	US-09-853-918-22
16	1856	70.9	384	10	US-09-853-918-21
17	1856	70.9	384	10	US-09-853-918-25
18	1856	70.9	384	10	US-09-853-918-27
19	1855.5	70.8	385	10	US-09-853-918-31

20	1855.5	70.8	385	10	US-09-853-918-33	Sequence 33, Appl
21	1855.5	70.8	385	10	US-09-853-918-35	Sequence 35, Appl
22	1855	70.8	384	10	US-09-853-918-23	Sequence 23, Appl
23	1853	70.8	384	10	US-09-853-918-28	Sequence 28, Appl
24	1852	70.7	384	10	US-09-853-918-26	Sequence 26, Appl
25	1851.5	70.7	385	10	US-09-853-918-39	Sequence 39, Appl
26	1644	62.8	384	10	US-09-853-918-30	Sequence 30, Appl
27	1210.5	46.2	371	9	US-10-076-157-9	Sequence 9, Appl
28	875	33.4	489	10	US-09-815-242-11953	Sequence 11953, A
29	874.5	33.4	217	10	US-09-925-302-829	Sequence 828, App
30	863	33.0	506	9	US-09-738-626-4172	Sequence 4172, App
31	856	31.9	488	10	US-09-815-242-10986	Sequence 10986, A
32	521	19.9	133	10	US-09-853-918-61	Sequence 61, Appl
33	492	18.8	477	9	US-09-738-626-6445	Sequence 6445, Ap
34	425.5	16.2	156	10	US-09-734-017A-76	Sequence 76, Appl
35	407	15.5	151	9	US-10-076-157-8	Sequence 8, Appl
36	317	12.1	325	10	US-09-815-242-12738	Sequence 12738, A
37	317	12.1	325	10	US-09-815-242-13035	Sequence 13035, A
38	316	12.1	325	10	US-09-815-242-10811	Sequence 10811, A
39	316	12.1	327	10	US-09-815-242-4895	Sequence 4895, Ap
40	312	11.9	325	10	US-09-815-242-11543	Sequence 11543, A
41	303.5	11.6	347	10	US-09-815-242-10035	Sequence 10035, A
42	302.5	11.6	347	10	US-09-815-242-13914	Sequence 13914, A
43	283	10.8	328	10	US-09-815-242-13261	Sequence 13261, A
44	173.5	6.6	265	10	US-09-925-300-1572	Sequence 1572, Ap
45	154	5.9	374	9	US-09-738-626-4174	Sequence 4174, Ap

## ALIGNMENTS

RESULT 1  
US-09-853-918-63  
Sequence 63, Application US/09853918  
Patent No. US20020068346A1  
GENERAL INFORMATION:  
APPLICANT: Kryslek, Stanley R.  
APPLICANT: Sherliff, Steven  
APPLICANT: Witmer, Mark R.  
APPLICANT: Hollenbaugh, Diane L.  
APPLICANT: Yan, Ning  
APPLICANT: Mouravieff, Julie E.  
APPLICANT: Einspahr, Howard M.  
APPLICANT: Kish, Kevin  
TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: DE24NP  
CURRENT FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 60/203,448  
PRIOR FILING DATE: 2000-05-10  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 63  
LENGTH: 514  
TYPE: PRT  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
AUTHORS: Collart, Frank R.  
AUTHORS: Huberman, Eliezer  
TITLE: Cloning and Sequence Analysis of the Human and Chinese  
JOURNAL: J. Biol. Chem.  
VOLUME: 263  
ISSUE: 30  
PAGES: 15769-15772  
DATE: 1988-10-25  
US-09-853-918-63  
Query Match 99.5%; Score 2605; DB 10; Length 514;  
Best Local Similarity 99.6%; Pred. No. 2,3e+211;  
Matches 512; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MADYISGTSVVPDDGLTAOOLFNCGGGLTYNDELLIPGYIDPFAQDYDLTSALTKTT	60
Db	1	MADYISGTSVVPDDGLTAOOLFNCGGGLTYNDELLIPGYIDPFAQDYDLTSALTKTT	60
QY	61	LKTPLVSSPMDVTVEAGAAIAMAALNGSIGTFIHNCNTPPEQANEEVKKVKKYEQGFITDPVY	120
Db	61	LKTPLVSSPMDVTVEAGAAIAMAALNGSIGTFIHNCNTPPEQANEEVKKVKKYEQGFITDPVY	120
QY	121	LSPKDRVADYFEAKRARGFCGIPITDTGRMGRLYGISSRDIIDLKKEEHDCFLIEIMT	180
Db	121	LSPKDRVADYFEAKRARGFCGIPITDTGRMGRLYGISSRDIIDLKKEEHDCFLIEIMT	180
QY	181	KREDIVAPRSTTLKEANFIIDORSKKKGLPIYNEDDEVAIIARPDLDKKNRDYPLASKDA	240
Db	181	KREDIVAPRSTTLKEANFIIDORSKKKGLPIYNEDDEVAIIARPDLDKKNRDYPLASKDA	240
QY	241	KKOLLGGAIGTHEDDKYRLDLLAAGADVVDVLDSSQGNISFOIMKIKYIKDKYPNLOYI	300
Db	241	KKOLLGGAIGTHEDDKYRLDLLAAGADVVDVLDSSQGNISFOIMKIKYIKDKYPNLOYI	300
QY	301	GGNVYTAOAOKLIDAGVDALPVMGSSGICTIOBYLACGRPOATAVYKYVEYARFPGV	360
Db	301	GGNVYTAOAOKLIDAGVDALPVMGSSGICTIOBYLACGRPOATAVYKYVEYARFPGV	360
QY	361	VIADGIGIONVGHIAKALALGASTVMGSLAATTEAPGEYFPFSDGIRLKKRYGMGSLDAM	420
Db	361	VIADGIGIONVGHIAKALALGASTVMGSLAATTEAPGEYFPFSDGIRLKKRYGMGSLDAM	420
QY	421	DKHLSSONRPFSEAKIRVAGVSGAVODKGSIMHFVYPLINGIGHSCODIGAKSLTOVR	480
Db	421	DKHLSSONRPFSEAKIRVAGVSGAVODKGSIMHFVYPLINGIGHSCODIGAKSLTOVR	480
QY	481	AMMYSGELKFEKRTSSAQVEGCVHSLHSYERLRF 514	
Db	481	AMMYSGELKFEKRTSSAQVEGCVHSLHSYERLRF 514	
RESULT 2			
US-09-853-918-49			
Sequence 49, Application US/09853918			
Patent No. US20020068346A1			
GENERAL INFORMATION:			
APPLICANT: Krysstex, Stanley R.			
APPLICANT: Sherliff, Steven			
APPLICANT: Wilmer, Mark R.			
APPLICANT: Hollenbaugh, Diane L.			
APPLICANT: Yan, Ning			
APPLICANT: Mouravieff, Julie E.			
APPLICANT: Einspahr, Howard M.			
APPLICANT: Kish, Kevin			
TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE			
TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF			
FILE REFERENCE: DB24NP			
CURRENT APPLICATION NUMBER: US/09/853,918			
PRIOR APPLICATION NUMBER: 60/203,448			
PRIOR FILING DATE: 2000-05-10			
NUMBER OF SEQ ID NOS: 65			
SOFTWARE: SeqIdN Ver. 2.0			
SEQ ID NO 49			
LENGTH: 514			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-09-853-918-49			

[illegible]

Qy	61	LKPLVSSPMDTVLEAGMAMTALGTGICFIHNHCSTPEQANEVRYKKKYEGFITDPVY	120
Db	61	LKTPLVSSPMDTVLEAGMAMTALGTGICFIHNHCSTPEQANEVRYKKKYEGFITDPVY	120
Qy	121	LSPKRDVDFEAKARHGFSGIPTTDGRMSRLVGIISRRIDFLKEEHDFLEIMT	180
Db	121	LSPKRDVDFEAKARHGFSGIPTTDGRMSRLVGIISRRIDFLKEEHDFLEIMT	180
Qy	181	KREDLVAPRSITTLKEANELLORSKKGKLPVNEDELVAITARTDLKKRRDYPLASKDA	240
Db	181	KREDLVAPRSITTLKEANELLORSKKGKLPVNEDELVAITARTDLKKRRDYPLASKDA	240
Qy	241	KKOLLCSAIGTHEDDKRLDLLAQAGDVYVLLDSGNSIFQJNMTKYTKDYPPNLOYI	300
Db	241	KKOLLCSAIGTHEDDKRLDLLAQAGDVYVLLDSGNSIFQJNMTKYTKDYPPNLOYI	300
Qy	301	GGNVVTAQAARKNLIDAGYDALRVGMSSGSIICIIODEVLACGRPOATAVYKYVEYARREGV	360
Db	301	GGNVVTAQAARKNLIDAGYDALRVGMSSGSIICITPEVLACGRPOATAVYKYVEYARREGV	360
Qy	361	VIAOGLIONVGHIAKALALGASTYMMGSLLAATTAEARGEFYFSFGITFLKTRYRGMGSLDAM	420
Db	361	VIAOGLIONVGHIAKALALGASTYMMGSLLAATTAEARGEFYFSFGITFLKTRYRGMGSLDAM	420
Qy	421	DKHLSQNRUYFSEADKIKVAQGVSGAVODKDSIKHFPYUULAGIQHSCODIGAKSLTQVR	480
Db	421	DKHLSQNRUYFSEADKIKVAQGVSGAVODKDSIKHFPYUULAGIQHSCODIGAKSLTQVR	480
Qy	481	AMMYSGELKEFERKTTSSAODEGVSHSYEKRRLP	514
Db	481	AMMYSGELKEFERKTTSSAODEGVSHSYEKRRLP	514

```

1      RESULT 3
2      US-09-853-918-62
3      ; Sequence 62, Application US/09653918
4      ; Patent No. US20020068346A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Krystek, Stanley R.
7      ; APPLICANT: Sheriff, Steven
8      ; APPLICANT: Witter, Mark R.
9      ; APPLICANT: Hollenbaugh, Diane L.
10     ; APPLICANT: Yan, Ning
11     ; APPLICANT: Moutavieff, Julie E.
12     ; APPLICANT: Einspahr, Howard M.
13     ; APPLICANT: Kish, Kevin
14     ; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE
15     ; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
16     ; FILE REFERENCE: DB24NP
17     ; CURRENT APPLICATION NUMBER: US/09/853,918
18     ; CURRENT FILING DATE: 2001-05-10
19     ; PRIOR APPLICATION NUMBER: 60/203,448
20     ; PRIOR FILING DATE: 2000-05-10
21     ; NUMBER OF SEQ. ID NOS: 65
22     ; SOFTWARE: PatentIn Ver. 2.0
23     ; SEQ ID NO 62
24     ; LENGTH: 514
25     ; TYPE: PRT
26     ; ORGANISM: Homo sapiens
27     ; PUBLICATION INFORMATION:
28     ; AUTHORS: Gu, Jing Jin
29     ; AUTHORS: Szychala, Jozef
30     ; AUTHORS: Mitchell, Beverly S.
31     ; TITLE: Regulation of the Human Inosine Monophosphate
32     ; TITLE: Dehydrogenase Type I Gene
33     ; JOURNAL: J. Biol. Chem.
34     ; VOLUME: 272
35     ; ISSUE: 7
36     ; PAGES: 4458-4466
37     ; DATE: 1997-02-14
38     ; US-09-853-918-62

```

Query Match 85.38; Score 2234; DB 10; Length 514;



Best Local Similarity 83.58; Pred. No. 3,8e-180;  
Matches 429; Conservative 39; Mismatches 46; Indels 0; Gaps 0

QY	1	MADVLISGCTSYVDDCGLTAAQLNFCNCDSJLTYNDFLLPCYIDETPAQJDLTSLATFKIT	60
Db	1	MADVLISGCTSYVEDGLTAAQLFASADGLTYNDFLLPGCFIDFIADZDLTSLATFKIT	60
QY	61	LKTPLVSSPMDVTVEAGMAIAMALTGIGFIHNCSTPEFOANEYRKVLYEOGFIIDPVV	120
Db	61	LKTPLLISSPMDVTVEADMALAMALMGIGFIHNCSTPEFOANEYRKVLYEOGFIIDPVV	120
QY	121	LSPDRARDVEAKARHFGCIPITPDGRMGSRVLVGIISSRDIDELKEHDCLEIIMT	180
Db	121	LSPHIVGVDTLEAMRGFSGIPITTEGTMGSKVLGVITSRDIIDELKEHDTLLSVMT	180
QY	181	KREDLVVAPRSTLTKEANELLORSKCKRLPVNEDDELVAIARTDLKRNDRPVLASKDA	240
Db	181	PRILEVAPACVTLKEANEILORSKCKRLPIVNCDELVAIARTDLKRNDRPVLASKDS	240
QY	241	KKOLLGCAIGTCHEDDKYRIDLLOAQADVVDLSDSGNSLFPQINMIRIKDKYPLNLOVT	300
Db	241	OKOLLGCAAVSTRFEDDKYRIDLLLOAQADVVDLSDSGNSVNOYAMVHTRIKOKYPLHLOV	300
QY	301	GGNVVTTAAQAKNLIDACVDALRVGMGSGSICIIODEVLACGRPOATAYKRYEVARRECPV	360
Db	301	GGNVVTTAAQAKNLIDACVDTGLRVGMGSGSICITIGEVMACGPGQSTAYKYAEVARRECPV	360
QY	361	VIAOGGIQNVGHIAKALALCALGASTYMGSLAATTAEAFGEFFSNGITLKKYRGMGSDAM	420
Db	361	IIAOGGIQTVGHVAVKALALGASTYMGSLAATTAEAGEYFFSGVRYKRYRGMGSDAM	420
QY	421	DKHLSSONRYSEADKILKVAGVSGAAYODKGIIRKFPYTLIAGIOMSGDIDIGAKSLTOVR	480
Db	421	EKSSSGOKRYSECDKVKIAQGVSGSLQDKGSIQKFPYTLIAGIOHCGDIDIGARSLSVLR	480
QY	481	AMMTSGELKFEKRTSSAQVDEGVASHSTYERLFL514	
Db	481	SMMTSGELKFEKRTMSAQIEGVGHLSTYERLY514	

Query Match	84.8%	Score 2222	DB 10	Length 514
Best Local Similarity	83.1%	Pred. No. 3.9e-179		
Matches 427	Conservative 39	Mismatches 48	Indels 0	Gaps 0
1 MADYLSGGTSYVPDDGLTAOOLFNCGGGLYNDFLILPGYIDETADQVPLTSAATRKII 60				

Db	1	MADYLISGCTGYVEDGLTQAOLVFASADDLTYNDFLLPGFIDFIADEVDLTSALTRKIT	60
Qy	61	LKTLTVSSPMOTVTEAGMATMALTGIGFIHNHCSTEPQANEKRYKKYKEOGFTIDPVV	120
Db	61	LKTLPISSPMOTVTEADMAIMALMGIGFIHNHCSTEPQANEKRYKKYKEOGFTIDPVV	120
Qy	121	LSPKDRVADVEAKARHGFCIPITPDGRMGSRLVGIISXKDIDFLKEEHDFLEIMT	180
Db	121	LSPSHTVGDVLEAKMRHGFGSIPITETGTMSKRLVGIIVTSRIDFLKEDHTTLESVMT	180
Qy	181	KREDLVAPRSITLKEANEIIQSRKKKGLRVNDEDELVAIATDCLKNRDPLASKDA	240
Db	181	PRIELVAPRVCTLKEANEIIQSRKKKGLRVNDCDELVAIATDCLKNRDPLASKDS	240
Qy	241	KKOLLGCAAICTHEDDKRYLDLLAQAQGVVULDDSSGNSIFQITMIAKTYIKDKYPNLOVT	300
Db	241	OKOLLGCAAVGTREDDKRYLDLLTQAQGVVULDDSSGNSVYQIYAMVHYIKOKYPHLOV	300
Qy	301	GGNVYTAQAQANLIDACVVDLRLVGMSSGSTCTIOEVLACGRQAATAYKKYKVEAARRGVP	360
Db	301	GGNVYTAQAQANLIDAGVGLRLVGMSSGSLCTIOEVLACGPGQATAYKKAEEAARRGVP	360
Qy	361	VIADGGIDONVGHIAKALALGASTYMMSSLLAATTABEGEYFFSDGIRLKYKRGMSLDAM	420
Db	361	IIADGGIDOTVGHVAKALALGASTYMMSSLLAATTABEGEYFFSDGVRLLKYKRGMSLDAM	420
Qy	421	DKHLISNORRYESMDKIRKVAQSGGAUODKSIHKFPYTLIAGIOHSCODIGAKSLTOVR	480
Db	421	EKSSSSQKRYTESDDKKYKIAQGVSGSTIODKCSIQKFPYTLIAGIOHSCODIGARSLSVLR	480
Qy	481	AMATYSGELKEFEKRTSSAONEGVGSHLSYEXERLTF	514
Db	481	SMATYSGELKEFEKRTMSAONEGVGSHLSYEXERLTV	514

: TITLE: Effects of Human T Lymphocyte Activation on Inosine  
 :  
 : TITLE: Monophosphate Dehydrogenase Expression  
 : JOURNAL: J. Immunol.  
 :  
 : VOLUME: 152  
 :  
 : PAGES: 984-991  
 :  
 : DATE: 1994  
 :

US-09-853-918-64

Query Match 84.1%; Score 2202; DB 10; Length 514;  
Best Local Similarity 82.1%; Pred No. 1.9e-177;  
Matches 422; Conservative 41; Mismatches 51; Indels 0; Gaps 0;

```
QY 1 MADYLLISGGSYVDDGLTAQOOLFNCGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MADYLLISGGYGVPEDEGLTAHELFAASADGLTYNDFLILPGYIDFTADEVDLTSLTRKIT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 LKPLVSSPMDITYTEAGMATAAMLTGIGIFIHNCIPEFOANEVRKYKKEGCFITDPVY 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 LKPLVSSPMDITYTEADMAIAMLTMGICIFIHNCIPEFOANEVRKYKKEGCFITDPVY 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 LSPKRDVDFEAKARHFGCGIPITDTGRMGSRLVGIISRDIDFLKEEHDFLEEIMT 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 LSPSHVGVDFLEAKMRHFGSGIPITETGTMGSKLVGIVTSRDIDFLAEKHDTLLSEVMT 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 KREDLVVAPRSITLKEANEILORSKKGKLPYVNEDELVAAIARTDLKKNRDYPLASKDA 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 PRIELVVPAGVTLKEANEILORTKKKGLPIVNDCELVAAIARTDLKKNRDYPLASKDS 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 KROLGGAATGTHEDDKRYLDLLAQAQGVVYVLDSSQNSIFQIMIKYIKDKYPIILOVY 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 OKOLGGAATGTHEDDKRYLDLLTQAQGVVYVPHSSQNSVYOIAMVHIKOKYPIILOVY 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 GGNVVTAAQAKNLLDAGVDALRVGMGSGSICIIOEVLACGRPOATAVYKVEYARREGVP 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 GGNVVTAAQAKNLLDAGVDGLRVGMCGSICITQEVWACGRPGGTAVYKAEYARREGVP 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 361 VIADGGIOWNGHIAKALALAGASTVMMGSLAATTEAPGEYFESDGLRLKKYRGMGSLDAM 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 361 IYADGGIOWNGHIAKALALAGASTVMMGSLAATTEAPGEYFESDGLRLKKYRGMGSLDPM 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 421 DKHLSSONRYFSEADRTKVAQVSGAVODKGSITHKFPYVLIAGIIOHSCODIGAKSLTOVR 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 421 EKSSSSOKRFFSGDKVKYIAQVSGSIOKGFVYPYLIAGIOHSCODIGARSLSVLR 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 481 AMMYSGELKEFKRTSSAOVEGGVHSLHSYEKRLP 514
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 481 SMYSGELKEFKRTMSPQIEGVGHLSHYEKRLY 514
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 6

US-09-853-918-65  
Sequence 65, Application US/09853918  
Patent No. US20020068346A1

GENERAL INFORMATION:

APPLICANT: Krystek, Stanley R.

APPLICANT: Sheriff, Steven

APPLICANT: Wilmer, Mark R.

APPLICANT: Hollenbaugh, Diane L.

APPLICANT: Yan, Ning

APPLICANT: Mouravieff, Julie E.

APPLICANT: Einspahr, Howard M.

APPLICANT: Kish, Kevin

TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE

FILE REFERENCE: DB24NP

CURRENT APPLICATION NUMBER: US/09/853,918

CURRENT FILING DATE: 2001-05-10

PRIOR APPLICATION NUMBER: 60/203,448

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 65

LENGTH: 514

TYPE: PRT

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

AUTHORS: Natsuneda, Yutaka

TITLE: Two distinct cDNAs for Human IMP Dehydrogenase

JOURNAL: J. Biol. Chem.

VOLUME: 265  
ISSUE: 9  
PAGES: 5292-5295  
DATE: 1990-03-25  
US-09-853-918-65

Query Match 84.0%; Score 2201; DB 10; Length 514;  
Best Local Similarity 82.3%; Pred. No. 2.3e-177;  
Matches 423; Conservative 39; Mismatches 52; Indels 0; Gaps 0;

```
QY 1 MADYLLISGGSYVDDGLTAQOOLFNCGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MADYLLISGGYGVPEDEGLTAQOOLFASADGLTYNDFLILPGYIDFTADEVDLTSLTRKIT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 LKPLVSSPMDITYTEAGMATAAMLTGIGIFIHNCIPEFOANEVRKYKKEGCFITDPVY 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 LKPLVSSPMDITYTEADMAIAMLTMGICIFIHNCIPEFOANEVRKYKKEGCFITDPVY 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 LSPKRDVDFEAKARHFGCGIPITDTGRMGSRLVGIISRDIDFLKEEHDFLEEIMT 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 LSPSHVGVDFLEAKMRHFGSGIPITETGTMGSKLVGIVTSRDIDFLAEKHDTLLSEVMT 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 KREDLVVAPRSITLKEANEILORSKKGKLPYVNEDELVAAIARTDLKKNRDYPLASKDA 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 PRIELVVPAGVTLKEANEILORSKKGKLPYVNDCELVAAIARTDLKKNRDYPLASKDS 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 KROLGGAATGTHEDDKRYLDLLAQAQGVVYVLDSSQNSIFQIMIKYIKDKYPIILOVY 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 OKOLGGAATGTHEDDKRYLDLLTQAQGVVYVPHSSQNSVYOIAMVHIKOKYPIILOVY 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 GGNVVTAAQAKNLLDAGVDALRVGMGSGSICIIOEVLACGRPOATAVYKVEYARREGVP 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 GGNVVTAAQAKNLLDAGVDGLRVGMCGSICITQEVWACGRPGGTAVYKAEYARREGVP 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 361 VIADGGIOWNGHIAKALALAGASTVMMGSLAATTEAPGEYFESDGLRLKKYRGMGSLDAM 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 361 IYADGGIOWNGHIAKALALAGASTVMMGSLAATTEAPGEYFESDGLRLKKYRGMGSLDPM 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 421 DKHLSSONRYFSEADRTKVAQVSGAVODKGSITHKFPYVLIAGIIOHSCODIGAKSLTOVR 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 421 EKSSSSOKRFFSGDKVKYIAQVSGSIOKGFVYPYLIAGIOHSCODIGARSLSVLR 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 481 AMMYSGELKEFKRTSSAOVEGGVHSLHSYEKRLP 514
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 481 SMYSGELKEFKRTMSPQIEGVGHLSHYEKRLY 514
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 7

US-09-853-918-20  
Sequence 20, Application US/09853918  
Patent No. US20020068346A1

GENERAL INFORMATION:

APPLICANT: Krystek, Stanley R.

APPLICANT: Sheriff, Steven

APPLICANT: Wilmer, Mark R.

APPLICANT: Hollenbaugh, Diane L.

APPLICANT: Yan, Ning

APPLICANT: Mouravieff, Julie E.

APPLICANT: Einspahr, Howard M.

APPLICANT: Kish, Kevin

TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE

FILE REFERENCE: DB24NP

CURRENT APPLICATION NUMBER: US/09/853,918

CURRENT FILING DATE: 2001-05-10

PRIOR APPLICATION NUMBER: 60/203,448

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 20

LENGTH: 384

TYPE: PRT

ORGANISM: Homo sapiens

Best Local Similarity 74.3%; Pred. No. 5,5e-149;

\_\_\_\_\_

Db 1 MADYLISGGTSYVPDDGLTAQOOLFNCGGDLTYNDLILPGYIDFTADQVDTLSALTKKIT 60  
QY 1 LKTPLVSSPMDVTVEAGMAIAMALTGIGIFIHNNCTPEFOANEVRKVKYEDGFTIDPVV 120  
Db 61 LKTPLVSSPMDVTVEAGMAIAMALTGIGIFIHNNCTPEFOANEVRKVKYQ----- 111  
QY 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVGISSRIDFLKEEHDFLEEIMT 180  
Db 112 ----- 111  
QY 181 KREDLVVAPRSITLKEANEILQSRKKGKLPVNEDELVAILIARTDLKKNRDYPLASKDA 240  
Db 112 -----POS----- 114  
QY 241 KKOLLGGAALIGHEDDKYRLDLAQAQVVDVVLDSQGSNIFQINMKIKYIKDKYPNLOYI 300  
Db 115 ---LLCGAALIGHEDDKYRLDLAQAQVVDVVLDSQGSNIFQINMKIKYIKDKYPNLOYI 171  
QY 301 GGNVYTAQAQAKNLIDAGVDALRVGMGSGSICIIQEVYLAGCRPOATAVYKVEYARFRGVP 360  
Db 172 GGNVYTAQAQAKNLIDAGVDALRVGMGSGSICIIQEVYLAGCRPOATAVYKVEYARFRGVP 231  
QY 361 VIADGGIQNVGHIKAKALALGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDM 420  
Db 232 VIADGGIQNVGHIKAKALALGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDM 291  
QY 421 DKHLSSQNRYPSEADKIKVAQVSGAVQDKGSIHKFPVPLIAGIQHSCODIGAKSLTOVR 480  
Db 292 DKHLSSQNRYPSEADKIKVAQVSGAVQDKGSIHKFPVPLIAGIQHSCODIGAKSLTOVR 351  
QY 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEKRLF 514  
Db 352 AMMYSGELKFEKRTSSAQVEGVHSLHSEKRLF 385  
RESULT 10  
US-09-853-918-29  
; Sequence 29, Application US/09853918  
; Patent No. US20020068346A1  
; GENERAL INFORMATION:  
; APPLICANT: Krystek, Stanley R.  
; APPLICANT: Sherliff, Steven  
; APPLICANT: Wilmer, Mark R.  
; APPLICANT: Hollenbaugh, Diane L.  
; APPLICANT: Yan, Ning  
; APPLICANT: Mouravieff, Julie E.  
; APPLICANT: Einspahr, Howard M.  
; APPLICANT: Kish, Kevin  
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
; FILE REFERENCE: POLYPEPTIDES AND USES THEREOF  
; CURRENT APPLICATION NUMBER: US/09/853,918  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 60/203,448  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-918-29  
Query Match 71.0%; Score 1860; DB 10; Length 384;  
Best Local Similarity 73.9%; Pred. No. 7,4e-149;  
Matches 380; Conservative 0; Mismatches 4; Indels 130; Gaps 1;

Db 61 LKTPLVSSPMDVTVEAGMAIAMALTGIGIFIHNNCTPEFOANEVRKVKYGGSG----- 113  
QY 121 LSPKDRVDVFEAKARHGCGIPITDTGRMGSRVGISSRIDFLKEEHDFLEEIMT 180  
Db 114 ----- 113  
QY 181 KREDLVVAPRSITLKEANEILQSRKKGKLPVNEDELVAILIARTDLKKNRDYPLASKDA 240  
Db 114 ----- 113  
QY 241 KKOLLGGAALIGHEDDKYRLDLAQAQVVDVVLDSQGSNIFQINMKIKYIKDKYPNLOYI 300  
Db 114 ---LLCGAALIGHEDDKYRLDLAQAQVVDVVLDSQGSNIFQINMKIKYIKDKYPNLOYI 170  
QY 301 GGNVYTAQAQAKNLIDAGVDALRVGMGSGSICIIQEVYLAGCRPOATAVYKVEYARFRGVP 360  
Db 171 GGNVYTAQAQAKNLIDAGVDALRVGMGSGSICIIQEVYLAGCRPOATAVYKVEYARFRGVP 230  
QY 361 VIADGGIQNVGHIKAKALALGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDM 420  
Db 231 VIADGGIQNVGHIKAKALALGASTVMGSLLAATTEAPGEYFFSDGIRLKKYRGMGSLDM 290  
QY 421 DKHLSSQNRYPSEADKIKVAQVSGAVQDKGSIHKFPVPLIAGIQHSCODIGAKSLTOVR 480  
Db 291 DKHLSSQNRYPSEADKIKVAQVSGAVQDKGSIHKFPVPLIAGIQHSCODIGAKSLTOVR 350  
QY 481 AMMYSGELKFEKRTSSAQVEGVHSLHSEKRLF 514  
Db 351 AMMYSGELKFEKRTSSAQVEGVHSLHSEKRLF 384  
RESULT 11  
US-09-853-918-36  
; Sequence 36, Application US/09853918  
; Patent No. US20020068346A1  
; GENERAL INFORMATION:  
; APPLICANT: Krystek, Stanley R.  
; APPLICANT: Sherliff, Steven  
; APPLICANT: Wilmer, Mark R.  
; APPLICANT: Hollenbaugh, Diane L.  
; APPLICANT: Yan, Ning  
; APPLICANT: Mouravieff, Julie E.  
; APPLICANT: Einspahr, Howard M.  
; APPLICANT: Kish, Kevin  
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
; FILE REFERENCE: POLYPEPTIDES AND USES THEREOF  
; CURRENT APPLICATION NUMBER: US/09/853,918  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 60/203,448  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-918-36  
Query Match 71.0%; Score 1859.5; DB 10; Length 385;  
Best Local Similarity 73.9%; Pred. No. 8,2e-149;  
Matches 380; Conservative 1; Mismatches 4; Indels 129; Gaps 1;

Db	111	----	110	----	110
Oy	181	KREDLVARPRSTILKEANEILÖRSKKGKLPVINEDEDELVAIIAFTDLKKNDFPLASKDA	240	----	300
Db	111	----	111	----	111
Oy	241	KKOLLGGAALIGTHEDDKRYDLLLAQAGDVVVLDSOGNSIFJOINMIRKIKDKYPLNÖVI	300	----	300
Db	112	GRPLTCGAALICTHEDDKRIIDLILAQAGDVVVLDSOGNSIFJOINMIRKIKDKDKYPLNÖVI	171	----	171
Oy	301	GGNVYTAQAANLLIDAGVDALRVGMSSGICIDIEVLACGRPOATAYVYEFYARRGVP	366	----	366
Db	172	GGNVYTAQAANLLIDAGVDALRVGMSSGICITDIEVLACGRPOATAYVYEFYARRGVP	231	----	231
Oy	351	VADGIONVGHILKALALGASTVMMSILATTEABEGEYFFSDGIRIKYRGMSLIDAM	420	----	420
Db	232	VADGIONVGHILKALALGASTVMMSILATTEABEGEYFFSDGIRIKYRGMSLIDAM	291	----	291
Oy	421	DKHLSSONRYFSEADRIKVAQVSGAVODKSIHKFPYLLAGIQHSGDITGAKSLTÖVR	480	----	480
Db	292	DKHLSSONRYFSEADRIKVAQVSGAVODKSIHKFPYLLAGIQHSGDITGAKSLTÖVR	351	----	351
Oy	481	AMNYSSELKFEKRTSSAOVCEGVSHLSYEKRLF	514	----	514
Db	352	AMNYSSELKFEKRTSSAOVCEGVSHLSYEKRLF	385	----	385

RESULT 12  
 US-09-853-918-24  
 Sequence 24, Application US/098653918  
 Patent No. US20020068340A1  
 GENERAL INFORMATION:  
 APPLICANT: Kyrestek, Stanley R.  
 APPLICANT: Sheriff, Steven  
 APPLICANT: Witmer, Mark R.  
 APPLICANT: Hollenbaugh, Diane L.  
 APPLICANT: Mouravieff, Julie E.  
 APPLICANT: Einspahr, Howard M.  
 APPLICANT: Kish, Kevin  
 TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
 TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
 FILE REFERENCE: DB2ANP  
 CURRENT APPLICATION NUMBER: US/09/853,918  
 CURRENT FILING DATE: 2001-05-10  
 PRIOR APPLICATION NUMBER: 60/203,448  
 PRIOR FILING DATE: 2000-05-10  
 NUMBER OF SEQ ID NOS: 65  
 SOFTWARE: PatentIn Ver. 2.10  
 SEQ ID NO 24  
 LENGTH: 364  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-853-918-24

[illegible]

Db	112	-----p1-----	113
Qy	241	KOLLGAAICTHEDDKYRLDLLAAGVDVVLDDSSGNSIFQINMITYIKDKYPNLOVT	300
Db	114	---LLCGAAICTHEDDKYRLDLLAAGVDVVLDDSSGNSIFQINMITYIKDKYPNLOVT	170
Qy	301	GGNVVTAQAANLLIDAGVDALRVGMSSGSICTIDEVLACGRPATAYKYKYEVARRGCP	360
Db	121	GGNVVTAQAANLLIDAGVDALRVGMSSSICITIDEVLACGRPATAYKYKYEVARRGCP	230
Qy	361	VIADGIONVGHIAKALALGASTVMGSSLAATTEAEGEPFSSGILTKYRGMGSLDAM	420
Db	231	VIADGIONVGHIAALALGASTVMGSSLAATTEAEGEYFFSGILTKYRGMGSLDAM	290
Qy	421	DKHLSSQNRVSEADKIKVAGVSGAVODKCSIHKFPYVLIAGIQHSCODIGAKSLTOVR	480
Db	291	DKHLSSQNRVSEADKIKVAGVSGAVODKCSIHKFPYVLIAGIQHSCODIGAKSLTOVR	350
Qy	481	AMMYSGELKFEKRTSSAIVEGGVSHISYERRLF	514
Db	351	AMMYSGELKFEKRTSSAIVEGGVSHISYERRLF	384

```

1      RESULT 13
2      US-09/853-918-37
3      ? Sequence 37, Application US/09853918
4      ? Patent No. US20020068346A1
5      ? GENERAL INFORMATION:
6      ? APPLICANT: Krystek, Stanley R.
7      ? APPLICANT: Sheriff, Steven
8      ? APPLICANT: Wilmer, Mark R.
9      ? APPLICANT: Hollenbaugh, Diane L.
10     ? APPLICANT: Yan, Ning
11     ? APPLICANT: Montavieff, Julie E.
12     ? APPLICANT: Einspahr, Howard M.
13     ? APPLICANT: Kish, Kevin
14     ? TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE
15     ? TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
16     ? FILE REFERENCE: DE2ANP
17     ? CURRENT APPLICATION NUMBER: US/09/853,918
18     ? CURRENT FILING DATE: 2001-05-10
19     ? PRIOR APPLICATION NUMBER: 60/203,448
20     ? PRIOR FILING DATE: 2000-05-10
21     ? NUMBER OF SEQ ID NOS: 65
22     ? SOFTWARE: PatentIn Ver. 2.0
23     ? SEQ ID NO 37
24     ? LENGTH: 385
25     ? TYPE: PRT
26     ? ORGANISM: Homo sapiens
27     ? US-09-853-918-37

```

Query Match	70.9%	Score 1857.5	DB 10	Length 385
Best Local Similarity	74.1%	Pred. No. 1.2e+18		
Matches 381	Conservative 0	Mismatches 4	Indels 129	Gaps 2
QY	1	MADYILSGSTSVPPDGLTAOOLFNCGGGATVINDSLITPGYIDFTADVDLTSLATKKT	60	
Db	1	MADYILSGSTSVPPDGLTAOOLFNCGGGATVINDSLITPGYIDFTADVDLTSLATKKT	60	
QY	61	LKTPVSSPMCTVTEAGMAIAMALTGGIGFTIHNCSTPEFOANEVRKVKYEGGFTDPV	120	
Db	61	LKTPVSSPMCTVTEAGMAIAMALTGGIGFTIHNCSTPEFOANEVRKVKY	110	
QY	121	LSFKORVRDVEFAKARHGCGIPITDGTGMSRLVGIISSRIDPLKEEHDCFLFELMT	180	
Db	111		110	
QY	181	KREDLVAPRSITLKEANEILDORSKKGLPTIVNEDDELVAIARTDLKKNRDYPLASKDA	240	
Db	111		114	
QY	241	KKOLLGGAIGTHEDDKYRLDLAAGVDVVLDSQGSNIFQIMIKYIKDKYPNLOVI	300	

Db 115 ---LLCGAATGTHEDDKYRLDLAQAQVDVYVLDSSQSGNSIFQINMKYIKDKYPNLQYI 171  
QY 301 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 360  
Db 172 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 231  
QY 361 VIADGCIQNVGHIAKALALAGASTVMGSLAATTEAPGEFFSDGIRLKKYRGMSLDM 420  
Db 232 VIADGCIQNVGHIAKALALAGASTVMGSLAATTEAPGEFFSDGIRLKKYRGMSLDM 291  
QY 421 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYLGIQHSODIGAKSLTOVR 480  
Db 292 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYLGIQHSODIGAKSLTOVR 351  
QY 481 AMYSGELKEKERTSSAOVEGVSHSYEKRLEF 514  
Db 352 AMYSGELKEKERTSSAOVEGVSHSYEKRLEF 385

## RESULT 14

US-09-853-918-38  
; Sequence 38, Application US/09853918  
; Patent No. US20020068346A1  
; GENERAL INFORMATION:  
; APPLICANT: Krystek, Stanley R.  
; APPLICANT: Sheriff, Steven  
; APPLICANT: Witmer, Mark R.  
; APPLICANT: Hollenbaugh, Diane L.  
; APPLICANT: Yan, Ning  
; APPLICANT: Mouravieff, Julie E.  
; APPLICANT: Einspahr, Howard M.  
; APPLICANT: Kish, Kevin  
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: DB24NP  
; CURRENT APPLICATION NUMBER: US/09/853,918  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 60/203,448  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-918-38

Query Match 70.9%; Score 1857.5; DB 10; Length 385;  
Best local Similarity 74.1%; Pred. No. 1.2e-148;  
Matches 381; Conservative 0; Mismatches 4; Indels 129; Gaps 2;

QY 1 MARYLLISGCTSYVPDDGLTFAOQLFNCGGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60  
Db 1 MARYLLISGCTSYVPDDGLTFAOQLFNCGGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60  
QY 61 LKTPPLVSSPMDTYTEGMAIAMAALTGIGIFIHNCCTPEFOANEVRKVKYEGGFITDPVY 120  
Db 61 LKTPPLVSSPMDTYTEGMAIAMAALTGIGIFIHNCCTPEFOANEVRKVKYEGGFITDPVY 112  
QY 121 LSPKDRYRVFEAKARHGCGIPIITDTGRMGRVLGIISSRIDIDFLKEEHDFLEEIIMT 180  
Db 113 -----PL----- 114  
QY 241 KROLLCGAATGTHEDDKYRLDLAQAQVDVYVLDSSQSGNSIFQINMKYIKDKYPNLQYI 300  
Db 115 ---LLCGAATGTHEDDKYRLDLAQAQVDVYVLDSSQSGNSIFQINMKYIKDKYPNLQYI 171  
QY 301 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 360

Db 172 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 231  
QY 361 VIADGCIQNVGHIAKALALAGASTVMGSLAATTEAPGEFFSDGIRLKKYRGMSLDM 420  
Db 232 VIADGCIQNVGHIAKALALAGASTVMGSLAATTEAPGEFFSDGIRLKKYRGMSLDM 291  
QY 421 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYLGIQHSODIGAKSLTOVR 480  
Db 292 DKHLSSQNRFFSEADKIKVAQGVSGAVQDKGSIHKFVPLYLGIQHSODIGAKSLTOVR 351  
QY 481 AMYSGELKEKERTSSAOVEGVSHSYEKRLEF 514  
Db 352 AMYSGELKEKERTSSAOVEGVSHSYEKRLEF 385

## RESULT 15

US-09-853-918-22  
; Sequence 22, Application US/09853918  
; Patent No. US20020068346A1  
; GENERAL INFORMATION:  
; APPLICANT: Krystek, Stanley R.  
; APPLICANT: Sheriff, Steven  
; APPLICANT: Witmer, Mark R.  
; APPLICANT: Hollenbaugh, Diane L.  
; APPLICANT: Yan, Ning  
; APPLICANT: Mouravieff, Julie E.  
; APPLICANT: Einspahr, Howard M.  
; APPLICANT: Kish, Kevin  
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: DB24NP  
; CURRENT APPLICATION NUMBER: US/09/853,918  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 60/203,448  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-918-22

Query Match 70.9%; Score 1857; DB 10; Length 384;  
Best local Similarity 74.1%; Pred. No. 1.3e-148;  
Matches 381; Conservative 0; Mismatches 3; Indels 130; Gaps 2;

QY 1 MARYLLISGCTSYVPDDGLTFAOQLFNCGGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60  
Db 1 MARYLLISGCTSYVPDDGLTFAOQLFNCGGDGLTYNDFLILPGYIDFTADQVDTLSALTKKIT 60  
QY 61 LKTPPLVSSPMDTYTEGMAIAMAALTGIGIFIHNCCTPEFOANEVRKVKYEGGFITDPVY 120  
Db 61 LKTPPLVSSPMDTYTEGMAIAMAALTGIGIFIHNCCTPEFOANEVRKVKYEGGFITDPVY 110  
QY 121 LSPKDRYRVFEAKARHGCGIPIITDTGRMGRVLGIISSRIDIDFLKEEHDFLEEIIMT 180  
Db 111 -SP----- 112  
QY 181 KREDLVVAPRSTILKEANEILQRSKKGLPIVNEDELVAIIARTDLKKNRDYPLASKDA 240  
Db 113 ----- 112  
QY 241 KROLLCGAATGTHEDDKYRLDLAQAQVDVYVLDSSQSGNSIFQINMKYIKDKYPNLQYI 300  
Db 113 ---SLCGAATGTHEDDKYRLDLAQAQVDVYVLDSSQSGNSIFQINMKYIKDKYPNLQYI 170  
QY 301 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 360  
Db 171 GGNVYTAQAQKNLIDAGVALRVGMGSGSICITQEVLAACGRPOATAVYKVEYARREFGP 230  
QY 361 VIADGCIQNVGHIAKALALAGASTVMGSLAATTEAPGEFFSDGIRLKKYRGMSLDM 420

Db	231	V I A D G I O N V G H I A K A L A L G A S T V M M G S L I A T T E A P G E Y F F S D G I R L K K Y R C M G S L D A M	290
Qy	421	D K H L S S O N R Y F S E A D K I K V A O G V S G A V O D K G S I H K F V P Y L I A G I O H S C O D I G A K S L T O V R	480
Db	291	D K H L S S O N R Y F S E A D K I K V A O G V S G A V O D K G S I H K F V P Y L I A G I O H S C O D I G A K S L T O V R	350
Qy	481	A M M Y S G E L K F E K R T S S A O V E G V H S L H S Y E K R L F	514
Db	351	A M M Y S G E L K F E K R T S S A O V E G V H S L H S Y E K R L F	384

Search completed: February 13, 2003, 11:17:22  
Job time : 306 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2003, 03:27:06 / Search time 98 seconds  
(without alignments)  
1080.696 Million cell updates/sec

Title: US-09-846-637C-4

Perfect score: 2619  
Sequence: 1 MADYLLISGGTSTVPDDGLTA.....SSAQEGVGHLSHSYKRLFF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPTREMBL\_21:\*  
2: sp.\_archaea:\*  
3: sp.\_bacteria:\*  
4: sp.\_fungi:\*  
5: sp.\_human:\*  
6: sp.\_invertebrate:\*  
7: sp.\_mammal:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_rodent:\*  
12: sp.\_virus:\*  
13: sp.\_vertebrate:\*  
14: sp.\_unclassified:\*  
15: sp.\_virus:\*  
16: sp.\_bacteriap:\*  
17: sp.\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2561	97.8	514	11 091211	091211 mus musculu
2	2182	83.3	445	11 09DCU6	09DCU6 mus musculu
3	2079.5	79.4	489	4 096NU2	096NU2 homo sapien
4	1597.5	61.0	521	3 09PBJ2	09PBJ2 candida alb
5	1563	59.7	529	3 09UVLO	09UVLO pneumocysti
6	1553.5	59.3	445	5 08SXMS	08SXMS drosophila
7	1416	54.1	534	5 09GZM3	09GZM3 caenorinabdi
8	1173.5	44.8	510	5 096387	096387 plasmodium
9	1148	43.8	502	10 092PA0	092PA0 glycine max
10	1117	42.6	501	10 09AY75	09AY75 oryza sativ
11	973	37.2	484	16 08RC64	08RC64 thermobacter
12	966.5	36.9	485	16 09PNN3	09PNN3 campylobact
13	958	36.6	485	16 097FM8	097FM8 clostridium
14	941.5	35.9	487	16 08XZG6	08XZG6 ralsstonia s
15	922	35.2	485	16 09KGN8	09KGN8 bacillus ha
16	914	34.9	487	16 09JUDO	09JUDO neisseria m

17	911	34.6	487	16 09JZB5	09JZB5 neisseria m
18	910.5	34.6	484	16 08X145	08X145 clostridium
19	910	34.7	485	16 09PAR5	09PAR5 xyella fias
20	904.5	34.5	482	16 09X168	09X168 thermotoga
21	901	34.4	487	16 08RE86	08RE86 fusobacteri
22	885	33.6	509	2 09RHG1	09RHG1 bacillus ce
23	878.5	33.5	509	2 09RHG1	09RHG1 bacillus ce
24	875	33.4	489	16 09HXMS	09HXMS pseudomonas
25	860	32.6	488	16 099W19	099W19 staphylococ
26	858.5	32.6	499	16 08YBK4	08YBK4 bruceella me
27	855.5	32.7	500	16 09RT87	09RT87 deinetococcus
28	848	32.4	392	10 0944T1	0944T1 glycine max
29	846.5	32.3	488	16 0926V9	0926V9 listeria mo
30	846	32.3	489	16 09KTW3	09KTW3 vibrio chol
31	836	31.9	500	16 09ZRT5	09ZRT5 rhizobium m
32	832.5	31.6	493	16 09C1Y6	09C1Y6 lactococcus
33	827	31.6	487	16 08ZCU3	08ZCU3 yersinia pe
34	825.5	31.5	506	2 09RH20	09RH20 coirnebacte
35	819.5	31.5	487	16 09A7V2	09A7V2 callobacter
36	818	31.2	488	16 08ZNS9	08ZNS9 salmonella
37	818	31.2	490	16 08Z4Q2	08Z4Q2 salmonella
38	815.5	31.1	503	16 08UH06	08UH06 agrobacteri
39	812.5	31.0	501	16 09L017	09L017 streptomyce
40	807.5	30.8	492	16 097NA3	097NA3 streptococc
41	793	30.3	500	16 0983F6	0983F6 rhizobium l
42	774.5	29.6	485	17 0978L4	0978L4 thermoplasm
43	747.5	28.5	485	17 09HLK8	09HLK8 methanopyru
44	746	28.5	502	17 08YV01	08YV01 methanobact
45	737.5	28.2	484	17 026245	026245 methanobact

## ALIGNMENTS

RESULT 1  
ID 091211 PRELIMINARY; PRT; 514 AA.

AC 091211:  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Inosine 5'-phosphate dehydrogenase 2.  
GN IMPDH2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL: BC010314; AAH10314.1;  
DR MGD: MGI:109367; Impdh2.  
DR InterPro: IPR000644; CBS domain.  
DR InterPro: IPR001093; IMPDH/GMPPrase.  
DR Pfam: PF00571; CBS; 2.  
DR Pfam: PF00478; IMPDH\_C; 1.  
DR Pfam: PF01574; IMPDH\_N; 1.  
DR TIGRfams: TIGR01302; IMP\_dehydrocog; 1.  
DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; UNKNOWN.1.  
SO SEQUENCE 514 AA; 55815 MW; 17025A5C5EBC439 CRC64;

Query Match 97.8%; Score 2561; DB 11; Length 514;  
Best local Similarity 98.1%; Pred. No. 2.1e-170;  
Matches 504; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MADYLLISGGTSTVPDDGLTAQOLFNGGDTYNDPILPGYIDFTADQVLTSA LRKKIT 60  
DB 1 MADYLLISGGTSTVPDDGLTAQOLFNGGDTYNDPILPGYIDFTADQVLTSA LRKKIT 60  
QY 61 LKTPLVSPMDVTEAGMALATGTCIGFIHNCCTPEQANVRKVKYEGCFITDPVV 120  
|||||

```

Db      61 LKTPVSSPMDVTYTAGMAIAMLATGGIGFIHNNCTPEFOANEVRKVKYEGGFTTDPVV 120
Oy      121 LSPKDRVDFVEFKARHGCPIPTDTGRHGSRLVGISSRDIDFLKEEHCDFLEIMT 180
Db      121 LSPKDRVDFVEFKARHGCPIPTDTGRHGSRLVGISSRDIDFLKEEHCDFLEIMT 180
Oy      181 KREDLVNARSITLKEANEILQSRKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
Db      181 KREDLVNARSITLKEANEILQSRKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
Oy      241 KQOLCGAALIGHEDDKYRLDLAAGVGVVLDSSQGSIFQIMIKYIKDKYNLOYT 300
Db      241 KQOLCGAALIGHEDDKYRLDLAAGVGVVLDSSQGSIFQIMIKYIKDKYNLOYT 300
Oy      301 GGNVYTAQAQAKNLIDAGVALRVGMSSGICITQEVLAACGRPOATAVYKVEARREGV 360
Db      301 GGNVYTAQAQAKNLIDAGVALRVGMSSGICITQEVLAACGRPOATAVYKVEARREGV 360
Oy      361 VIADGSIQNVGHITAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
Db      361 VIADGSIQNVGHITAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAM 420
Oy      421 DKHLSSONRFFSADKIKYAQGVSAVODKSGIHKFPYPLAGIOSHSCODIGAKSLTOVR 480
Db      421 DKHLSSONRFFSADKIKYAQGVSAVODKSGIHKFPYPLAGIOSHSCODIGAKSLTOVR 480
Oy      481 AMYSGELFEKERTSSAOVEGGVHSLHSEKRLF 514
Db      481 AMYSGELFEKERTSSAOVEGGVHSLHSEKRLF 514

RESULT 2
O9DCL6 PRELIMINARY: PRT: 445 AA.
AC 09DCL6:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Inosine 5'-phosphate dehydrogenase 2.
GN IMPDH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217651;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderrelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK002676; BAB22278.1; -.
DR HSSP: P12268; 1B30.
DR MGD: MGI:109367; Impdh2.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.

```

```

DR InterPro: IPR001093; IMPDH/GMPase.
DR Pfam: PF00571; CBS_2.
DR Pfam: PF00478; IMPDH_C_1.
DR Pfam: PF01574; IMPDH_N_1.
DR SMART: SM00116; CBS_2.
DR TIGRfams: TIGR01302; IMP_dehydrog.1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
SQ SEQUENCE 445 AA; 48474 MW; 1B3A400640CB4C9 CRC64;

Query Match
Best Local Similarity 96.9%; Pred No. 4,7e-144;
Matches 431; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Oy      70 MDTVTAGMAIAMLATGGIGFIHNNCTPEFOANEVRKVKYEGGFTTDPVYLSPKDRVD 129
Db      1 MDTVTAGMAIAMLATGGIGFIHNNCTPEFOANEVRKVKYEGGFTTDPVYLSPKDRVD 60
Oy      130 VFEAKARHGCPIPTDTGRMGSRLVGISSRDIDFLKEEHCDFLEIMTKREDLVNAP 189
Db      61 VFEAKARHGCPIPTDTGRMGSRLVGISSRDIDFLKEEHCDFLEIMTKREDLVNAP 120
Oy      190 RSTTLKEANEILQSRKKGKLPYNEDELVAIARTDLKKNRDYPLASKDAKQLCGAA 249
Db      121 AGVTLEANEILQSRKKGKLPYNEDELVAIARTDLKKNRDYPLASKDAKQLCGAD 180
Oy      250 IGHEDDKYRLDLAAGVGVVLDSSQGSIFQIMIKYIKDKYNLOYTIGANVYTAQA 309
Db      181 IGHEDDKYRLDLAAGVGVVLDSSQGSIFQIMIKYIKDKYNLOYTIGANVYTAQA 240
Oy      310 AKNLIDAGVALRVGMSSGICITQEVLAACGRPOATAVYKVEARREGVYARPDAGGION 369
Db      241 AKNLIDAGVALRVGMSSGICITQEVLAACGRPOATAVYKVEARREGVYARPDAGGION 300
Oy      370 VGHITAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAMPKHLSSONR 429
Db      301 VGHITAKALALGASTVMGSLAATTEAPGEYFSDGIRLKKYRGMSLDAMPKHLSSONR 360
Oy      430 YFSEADKIKYAQGVSAVODKSGIHKFPYPLAGIOSHSCODIGAKSLTOVRAMYSGEIK 489
Db      361 YFSEADKIKYAQGVSAVODKSGIHKFPYPLAGIOSHSCODIGAKSLTOVRAMYSGEIK 420
Oy      490 FEKRTSSAOVEGGVHSLHSEKRLF 514
Db      421 FEKRTSSAOVEGGVHSLHSEKRLF 445

RESULT 3
O96NU2 PRELIMINARY: PRT: 489 AA.
AC O96NU2:
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CDNA FLJ30078 fis. clone BGC112000533. highly similar to
DE Inosine 5'-monophosphate dehydrogenase 2 (EC 1.1.1.205).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oshima A., Takahashi F., Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
RA Morikawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki T.,
RA Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK054640; BAB70780.1; -.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001295; DHO_dh.

```

Query Match	Best Local Similarity	Matches 327: Conservative	Score 1597.5: DB 3: Length 521: Pred. No. 3.6e-103: Mismatches 107: Indels 15: Gaps
EMBL: AF249293; AAF70813.1; -			
HSSP; P12268; 1B30.			
DR InterPro: IPR000644; CBS_domain.			
DR InterPro: IPR003009; FMN_enzyme.			
DR InterPro: IPR001093; IMPDH/GMPase.			
DR Pfam: PF00571; CBS; 2.			
DR Pfam: PF00478; IMPDH_C; 1.			
DR Pfam: PF01574; IMPDH_N; 1.			
DR SMART: SMO0116; CBS; 2.			
DR TIGRFAMs: TIGR01302; IMP_dehydrog; 1.			
DR PROSITE: PS00487; IMP_DH_GMP_REP; 1.			
FT VARIANT 47 47 V -> I.			
FT VARIANT 102 102 A -> S.			
FT VARIANT 400 400 D -> G.			
SEQUENCE 521 AA; 56267 MW; 7A1CF4DF6184FE7E CRC64;			
7 SGGTSVYPD---DGLTAQQLFNCGD--GLTYNDPLLPYIDFTADQVDTLSATLKRT 60	61.04; 63.18; 69; 107; 15; 103;		
6 SKATSYLDKYEFKKDLQSVKELIDSTNGTLVNDPELILPVLVPSASVLEKRLKIT 65			
QY 61 LKTPLVSSPMPTTYEAGMALTMALTGIGTGIHHNCIPPEAPANEVKKYKKEGCTITDPV 120			
66 LKSPVSSPMPTTYEENNAIMHALLGIGIHHNCIAEQAQEVKKYKKEGCTINDPV 125			
QY 121 LSPDRVADVEAKARHFGCGIPITDTGRMGSRVLVGISSRDIDFLKEEHDCLFEIMT 180			
126 ISPEVTAEVEYKMEGVLTGTSFPTENGKVGKGLVGIITSRDIOF--HEDNKSPEVMT 183			
QY 181 KREDLVYAPRSITLKEANELLORSKKKPLVNEDEELVIAARTLQKKNROYPLASKD 239			
184 K--DLVGGKKOISLTQGNELRSKKKPLVDAEGNLVSLISRTIDQKNODIPNASKSF 241			
QY 240 AKKQLLGCALIGTHEDDKYRLDLQAQGVVVVVLDSQSGNSIFQIMKIYITDKYVPLNV 299			
242 HSKQLLGCALIGTIDADRERLDKLVLEGVLVVVLDSGSSVQLMKIKIKREKYPELOV 301			
QY 300 IGGNVVYAAQAKNLLIDAGVDALRYMGSGSICIIQFLACGRQAPAVKYVEYARFQV 359			
302 IAGNVVTEQALLIEGADALRYMGSGSICITQEYMACGRQAGTAVVGVTEFAKRFV 361			
QY 360 PVIADGIGIYNGHIAKALALGASTVMGSLAATTEAPCEFFSDGIRLUKKYRGMSLDA 419			
362 PCIADGIGIGNCHITKALALGASVVMGSLAGLACTAEPDDYFPRDGRRLKTYRGMSIDA 421			
QY 420 MDKHLSSQN---RYSEADKIKYAAQGVSAVQADKSGIHFPVYLILAGIIGHSGODIGARS 475			
422 MQQNTNTNANSTSRFSEADVLVLAQGVSSVVDKSGITKRFVPLVYNGIHSYLDIDIGKS 481			
QY 476 LTQVRAAMYSGELKFEKRTSSAQVEGVSHSYEKRL 513			
Db 482 IDELEENVNNGEIRFERRTASAOEGGVGHLSHYEKRL 519			
RESULT 5			
OC 09UVI0			
ID 09UVI0	PREDIMINARY;	PRT;	529 AA.
AC 09UVI0;			
DT 01-MAY-2000 (TIREMBLrel. 13, Created)			
DT 01-MAY-2000 (TIREMBLrel. 13, Last sequence update)			
DT 01-JUN-2002 (TIREMBLrel. 21, Last annotation update)			
DE Inosine 5'-monophosphate dehydrogenase.			
OS Pneumocystis carinii.			
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae			
OC Pneumocystis			
OC NCBI_TaxID=4754;			
OC [1]			
RN SEQUENCE FROM N.A.			
RP MEDLINE=21152226; Pubmed=11223253;			
RA Ye D., Lee C.H., Queener S.F.;			

```
RT "Differential splicing of Pneumocystis carinii f. sp. carinii inosine
RT 5'-monophosphate dehydrogenase pre-mRNA."
RL Gene 263:151-158(2001).
DR EMBL: AF196975; AAF13230.1; -.
DR HSSP: P12268; 1B30.
DR InterPro: IPR000664; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPDH/GMPrtase.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF01574; IMPDH_N; 1.
DR SMART: SM00116; CBS; 2.
DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
SQ SEQUENCE 529 AA; 58066 MW; 030573A8854ADB5A CRC64;

Query Match 59.7%; Score 1563; DB 3; Length 529;
Best Local Similarity 61.4%; Pred. No. 9,5e-101;
Matches 316; Conservative 69; Mismatches 112; Indels 18; Gaps 7;

QY 12 YVPDDGLTAQQLFNCG---DGLTYNDFLLPGYIDFTADQVDLTSAITKTKITLPLVSS 68
DB 12 YSKDCGYDLDLSL-CRRIRGGLTYNDIILPGYIDFEVNSVLSHSITTKIYVLTPTFMS 78
QY 69 PMDTVEAGMAIMALTGGIGFTIHNCPTPEQANERYKVKYKKEQFTTDPVLSPKDRVR 128
DB 79 PMDTVESEMAIMALTGGIGFTIHNCPTPEQANERYKVKYKKEQFTTDPVLSLNRHVR 138
QY 129 DVEEAKARHFGCGIPITDGRMGSRVGISSRDIDELKEEHDCELEETMKREDLVYA 188
DB 139 DVARRIKEELGSPITDYGOLKGLGIVTSRDIOFHNND--SFLSEVTK--DLVVG 194
QY 189 PSITLKEANEILQSRKSKKLPIVNEDEDLVAIARTDLKKNRDPYLPASK-DARKOLLG 247
DB 195 SEGIRLEANEILRSCKKKGKPLVDKEGNLTALSRSDLMKNHFLASKLPDSKQLICA 254
QY 248 AALGTHEDKRYLDLQAQGVYVYLDSSGNSIFQINMKIKYKDPYLVQVIGVAVTA 307
DB 255 AAVGTGPPDRIIRKLHLYEGLDVLVDSSGNSIYOINMKIKKFFPNLEVTAGNVYTR 314
QY 308 AAKKNLIDAGVDLRYGMSGSCITIOFLACGRPOATAVYKVEYARFPGVAVDGGI 367
DB 315 EGMANLISGADLRVGMGSGSCITIOELMAVGRPOATAVAYSEASRFGVPTIADGGI 374
QY 368 QNVGTHAKALGASTVMGSLAATTEARGEYFFSDGIRLKKYRGMGSLDAMDKHLSSQ 427
DB 375 ENGHITTKALGASAVVMGNLLAGTTESPGYRRDQGRKLSYRGMGSLDAME-HLSGK 433
QY 428 N-----RYSEADKIKVAGVSGAVODKGSIHKFVYPLLGIQHSODIGAKSLTOY 479
DB 434 NKGDNAASSRYGFEADTIRVAGVSGVIDKGSILHAYVYPLRTGLQHSODIGVQNLTEL 493
QY 480 RAMVYSGELFEKRTSSAQVEGCVHLSHYEKRLF 514
DB 494 KQOVKEKNIRFERRYVASOLEGNVHGLDSYQKKLM 528

RESULT 6
OBSXMS PRELIMINARY: PRT; 446 AA.
AC OBSXMS:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)
LD36080p.
CN RAS.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY089553; AAL90291.1; -.
SQ SEQUENCE 446 AA; 48139 MW; 15746E0A799702BB CRC64;

Query Match 59.3%; Score 1553.5; DB 5; Length 446;
Best Local Similarity 67.5%; Pred. No. 3.4e-100;
Matches 303; Conservative 63; Mismatches 76; Indels 7; Gaps 3;

QY 70 PMDTVEAGMAIMALTGGIGFTIHNCPTPEQANERYKVKYKKEQFTTDPVLSPKDRVRD 129
DB 1 MOTVTESEMAIMALCGGIGIITHNCPTPEQALEVHKVKYKHGFMRDPSVMSPTTVGD 60
QY 130 VFEAKARHFGCGIPITDGRMGSRVGISSRDIDELKEEHDCELEETMKREDLVYAP 189
DB 61 VLEARRKNFGTGYPTENGKLGKGLGVTSRDIDE-RENOPEVLLADIMT--TELVTAP 117
QY 190 RSITLKEANEILQSRKSKKLPIVNEDEDLVAIARTDLKKNRDPYLPASKDARKOLLGCA 249
DB 118 NGINLPTANALIEKSKKGLPIVNGAGELVAMIARTDLKARSYPNASKNSKOLLVGA 177
QY 250 IGTBEDKRYLDLQAQGVYVYLDSSGNSIFQINMKIKYKDPYLVQVIGVAVTAQ 309
DB 178 IGTREDKARLALLVANGVDVLIIDSSGNSYOVEMIKYIKETPELOYIGGNVYTRQ 237
QY 310 AKNLIDAGVDALRVGMSGSCITIOFLACGRPOATAVYKVEYARFPGVPIADGION 369
DB 238 AKNLIDAGVDGLRVGMSGSCITIOEVMACGPOATAVYQVSTYARQFGVPIADGIGS 297
QY 370 VGHAKALGASTVMGSLAATTEARGEYFFSDGIRLKKYRGMGSLDAMD----KHLIS 425
DB 298 IGHVYALGASAVVMGSLAGTSEAPGEYFFSDGIRLKKYRGMGSLDAMERGAKGAA 357
QY 426 SQNRYPEADKIKVAGVSGAVODKGSIHKFVYPLLGIQHSODIGAKSLQVRAMYS 485
DB 358 MSRYINEMDKKRVAGVSGSIVDKGSVLRPLYLECGIQHSODIGANSINKLRMITYN 417
QY 486 GELKFEKRTSSAQVEGCVHLSHYEKRLF 514
DB 418 GOLRFMKRTHSQALEGNVHGLFSYEKRLF 446

RESULT 7
OBSXMS PRELIMINARY: PRT; 534 AA.
AC OBSXMS:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Hypothetical 58.2 kDa protein.
LD22D1.3.
CN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Geisel C., Bradshaw H., Hawkins M.;
```



```

RP SEQUENCE FROM N.A.
RA Cao Y., Schubert K.R.;
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AJ010201; CAB8030.1; -.
DR HSSP: P12268; 1B30.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPdh/GMPcase.
DR Pfam: PF00571; CBS; 1.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF01574; IMPDH_N; 1.
DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
SQ SEQUENCE 502 AA; 53407 MW; 617AB87613C36AA4 CRC64;

Query Match 43.8%; Score 1148; DB 10; Length 502;
Best Local Similarity 48.8%; Pred. No. 8e-72;
Matches 247; Conservative 78; Mismatches 163; Indels 18; Gaps 7;

QY 15 DDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTITKTLKPLVSSPMDTVT 74
DB 9 EDFFTAELKFTQGFSTTYDYVILPHRIDPAADAVDLSTLRPLAVFVASPMDTV 68

QY 75 EAGMAIAMLTGIGIFLHNCTPEFOANEVRYKYEQGFITDPVYLSPKDRVR--DVEE 132
DB 69 ESAMAAAMASLGIAVYHNSVPAAYQAAILRAKSRVRPILSDPAFAAPSAVEHDDAFG 128

QY 133 AKRHGFCGPIPTDTGMRGLVGIISRPIDFLKEEHCCFLEIMTKREDLVAPRSI 192
DB 129 ASP----FLLVYDTGTSGVRLGLGVARS--DWTNQDKLRYGDVYAPPPK--PAFWNA 179

QY 193 TLKEANELIORSKKGKLPYNEDEDLVAIARTDLKKNRDPYPL---ASKDAKKOLLGCA 248
DB 180 DLKKNIMESEKSGANAL-ERDGEVVDLVYREEVERRGTPKPLAATYAGAGEFVYGA 238

QY 249 AIGTHEDDKYRLDLAAGAVDVVVLDSQGSNIFQIMIKYIKDKYPNLOVIGSNVTA 308
DB 239 AVGTREDDKERLEHLVAGLVNVVLDSSQGSNIYOLEMNVYKRVYBELDIVIGSNVTMY 298

QY 309 QAKNLTADGADALRVGSGSICITIOEVLACGRPOATAVYKVEYARFPGVPIYADSGIQ 368
DB 299 QAKNLTADGADALRVGSGSICITIOEVLACGRPOATAVYKVEYARFPGVPIYADSGIS 358

QY 369 NVGHAKALALGASTYVMGSLAATTEARPEGEFFSDGIRLKKYRGMSLDAMDKHLSSQN 428
DB 359 NSGHYKALSLGASTYVMGSLAATTEARPEGEFFSDGIRLKKYRGMSLDAMDKHLSSQN 416

QY 429 RYTSSEADKIVAGVSGAVODKSIHKFVRYLLAGIOHSCODIGAKSLTOVRAMYSSEL 488
DB 417 RYLGDPYAKLIVAGVSGAVODKSVLNFIPYTLQAVRQGFODIGASSLSQAHDLRLREL 476

QY 489 KFEKRTSSAAGEGVHLSHYEKRLF 514
DB 477 KLEVRSGAAQVEGCVHGLVSEYKRYF 502

RESULT 10
Q9AY75 PRELIMINARY; PRT; 501 AA.
AC 09AY75;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 21, Last sequence update)
DE Putative inosine monophosphate dehydrogenase.
GN OSJNB0091J19.1.
OS Oryza sativa (rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OC NCBI_TaxID=4530;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. NIPPONBARE;
RC

```

```

RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zisman V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
RA Bowman C.L., Craven B., Uitterback T.R., Khalak H., Feldblum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
RT "Oryza sativa chromosome 3 BAC OSJNB0091J19 genomic sequence.";
RA EMBL: AC084320; AAK09225.1; -.
DR HSSP: P12268; 1B30.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPdh/GMPcase.
DR Pfam: PF00571; CBS; 1.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF01574; IMPDH_N; 1.
DR SMART: SM00116; CBS; 1.
DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
SQ SEQUENCE 501 AA; 52655 MW; DDE295CB714C3802 CRC64;

Query Match 42.6%; Score 1117; DB 10; Length 501;
Best Local Similarity 47.2%; Pred. No. 1.2e-69;
Matches 247; Conservative 80; Mismatches 160; Indels 36; Gaps 9;

QY 6 ISGGSYVDDGLTAQQLFNCGDGLTYNDLILPGYIDFTADQVDTLSALTITKTLKPL 65
DB 1 MAASSADLADGCPAPRLPSQGVSYTYDVIFLPGYIGFPADAVDLSTLRIRPLSIPC 60

QY 66 VSSPMDTVTFAGAAIMALTGIGIFLHNCTPEFOANEVRYKYEQGFITDPVYLSPMD 125
DB 61 VASPMPTVSSAAAMAAASLGAAYVCHNTEPHLOASIVKAAASRRPLPVSSVPLSPS 120

QY 126 --RVRVDFEAKRHGFCGPIPTDTGMRGLVGI-----ISSRPIDFLKEEHCCFLEI 178
DB 121 TPRLSDP---AGHDV-GL-VTERGDSLSKLVGVAAEHSR-----QAP 160

QY 179 MKREDLVAPRSIT---LKEANELIORSKKGKLPYNEDEDLVAIARTDLKKNRDP 234
DB 161 LPVSEYMRPAPRSVSASFDEQAALFADGLDYAPLVSDSEVIDLITVNDVERIRSP 220

QY 235 LASK---DAKKOLLGCAIGTHEDDKYRLDLAAGAVDVVVLDSQGSNIFQIMIKYIK 291
DB 221 KLCPSIGADGKFFVAAISITREDDKRLRLQVLKAGANALVSSQGSNIYQIMDKRYAK 280

QY 292 DKYPNLOVIGSNVTAQAQNLIDAGVDALRVGSGSICITIOEVLACGRPOATAVYKY 351
DB 281 KMYPEVDLIGSNVTVTAQAQNLIVAGVDGLRVGSGSICITIOEVLACGRPOATAVYKY 340

QY 352 EYARRGVPYIADGGIQONVGHITAKALALGASTYVMGSLAATTEARPEGEFFSDGIRLKY 411
DB 341 SYAKDHVNPYIADGGISNSGHIVKALSLGASTYVMGSLAATTEARPEGEFFSDGIRLKY 400

QY 412 RKGSLDAMDKHLSSQNRYSSEADKIVAGVSGAVODKSIHKFVRYLLAGIOHSCODI 471
DB 401 RKGSLDAMDKHLSSQNRYSSEADKIVAGVSGAVODKSVLRFPIYTMQAVKQGFOLD 458

QY 472 GAKSLTOVRAMYSSELKFEKRTSSAAGEGVHLSHYEKRLF 514
DB 459 GASSLSQAHDLRLRELSTIKLEVTGAAQVEGCVHGLVSEYKRAF 501

RESULT 11
Q8RC64 PRELIMINARY; PRT; 484 AA.
AC 08RC64;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE IMP dehydrogenase/GMP reductase.
GN GUAB OR TTE0582.
OS Thermoaerobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoaerobacteriales; Thermoaerobacteriaceae; Thermoaerobacter.
OX NCBI_TaxID=119072;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MB4T / JCM11007:
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W.,
  Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma
  Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL: AE013027; AAM23853.1; -
KM Complete proteome.
SQ SEQUENCE 484 AA: 52460 MW: 6531013B9BD78EA9 CRC64

Query Match          37.2%; Score 973; DB 16; Leng h 484;
Best Local Similarity 45.1%; Pred. No. 1.2e-59;
Matches 216; Conservative 89; Mismatches 160; Ind ls 14; Gaps 5;

OY 28 DCLTYNDPLPGYIDFTADVDLTSAITKTKITKPLVSSPMQVTE GMATAMALTGG 87
DB 8 EGLTFEDVLLIRAKSVLPKDYDLTKRLTKITLNIPLKSAGMDVTE RLAIATARECG 67
OY 88 IGFIIHNCIPEFOANEVRKVKRYEEOGFIITDPVYLSPKDRVDVFEAKA HGFCGIPITDT 147
DB 68 IGVIIHNMSEIHOAMEVDRKSEHGVIITDPFSLSPDHTIRDAELMA YKISGVPIIT-- 125
OY 148 GMSGRLVGIISRDIDFLKEEHDCFLFEIMTKRDLVVAARSTILK ANELORSKKG 207
DB 126 --VDSKLVGIITNRIDF--EDDLKPIREVMTK--ENLTAPPGTLE AKOILKKHKE 180
OY 208 KLPIVNEDELVAIIARTDLKKNRDYPLASKDAKKOLLGCAAICTHED KYRDLDLAOG 267
DB 181 KLPLVDENNVLGLITIKIEKAVEFPNAKDEKGLVLAAMAVCYCKD MDRVKALITEAG 240
OY 268 VVVVVLDSQGSITFOINMIKIKDKYPNLOYIGGNVNTAAOKNLID GVDALRYGMS 327
DB 241 VVAIVVDAHAGHSTRVLAIVAKIKKYPVOLIAGVNAFAEATRIDIE GADAVVKGIOP 300
OY 328 GSICTIOEVLACGRPOATVAVKYVEARFGVPVIADGGIOWNGHIAT LALGASVVMG 387
DB 303 GSICITRVAVAGVPOIATIAIYECAKEADKYGIPIVADSGIKISGDIYK IAGASVVMIG 360
OY 388 SLAATTEAPGEYFFSDGIRLKKYRGMSLDAMDKHLSSONRYFSEAD IKVAQGVGAV 447
DB 361 SLFACTEESPGIEIYOGSYKYVYRGMSLGAMKE--GSSDRYFOBEA KLVPEGVEGRV 418
OY 448 GOKGSIHKFVPLIAGIOHSCODIGAKSLTOVRAMYSGLKFEKRTS AOVEGGVHSL 506
DB 419 PVRGFLRETVIOLIGLGRAGMGYCGVKNIIELT--KTKFVRITQ GLIESHPDIT 472

RESULT 12
O9PNM3 PRELIMINARY: PRT: 485 AA.
AC O9PNM3:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
GN GUAB OR CJI058C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
  Jolley K., Kariyasev A.V., Mouton R., Pallen M.J., Pen
  Nall M.A., Rajadream M.A., Rutherford K.M., Van Vliet A.H.M.,
  Whitehead S., Barrett B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni"
```

```

RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL: AL139077; CAB73314.1; -
DR HSSP; P12268; 1B30.
DR InterPro: IPR000544; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPDH/GMPrtase.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF01574; IMPDH_N; 1.
DR SMART; SM00116; CBS; 2.
DR TIGRFAMs; TIGR01302; IMP_dehydrog; 1.
DR PROSITE; PS00487; IMP_DH_GMP_RED; 1.
KM Complete proteome.
SQ SEQUENCE 485 AA: 52151 MW: 63433266B93CDF CRC64;

Query Match          36.9%; Score 966.5; DB 16; Length 485;
Best Local Similarity 45.6%; Pred. No. 3.5e-59;
Matches 219; Conservative 78; Mismatches 164; Indels 19; Gaps 8;

OY 30 LTYNDFLLPGYIDFTADVDLTSAITKTKITKPLVSSPMQVTEAGMAIALTGGIG 89
DB 8 LTFEDVLLIRAPGSEVLPRKVKITHKTKITLIMPLISAMDTVEHRAIMMARLGGIG 67
OY 90 FIHNCIPEFOANEVRKVKRYEEOGFIITDPVYLSPKDRVDVFEAKRHGFCGIPITDTGR 149
DB 68 VHKMMDIASQVREVKRVKSSGVIIIDPIFVSPASVALEIMAEYRISGVPVADDEK 127
OY 150 MGSRLVGIISRDIDFLKEEHDCFLFEIMTKRDLVVAARSTILKANELIORSKKGK 209
DB 128 ---KLIIGITNRIDLF--ESDSNLVENVMTKM-PLITAPKCGCTLDIAEKIFSTKVEKL 181
OY 210 PIVNEDELVAIIARTDLKKNRDYPLASKDAKKOLLGCAAITGHEDDKYRDLDLAOGVD 269
DB 182 PIVDEGRLEGLEGIITDKKRREYDPDANKNGRLRGALIVGQMD--RVDAIVEAGVD 239
OY 270 VVVVVLDSQGSITFOINMIKIKDKYPNLOYIGGNVNTAAOKNLIDAGVDALRYGMS 329
DB 240 VVVLDSAHGHSGLIIDYKAIKAKIPNLDLITAGNATATAAMAKALCEAGDAVKVIGIOS 299
OY 330 ICIIOEVLACGRPOATVAVKYVEARFGVPVIADGGIOWNGHIKALALASTVMSGL 389
DB 300 ICTTRIVSGVGPQISADECEVEAFANKRGVPVIADGGIKYSGSDIAKALAVGASVMICSL 359
OY 390 LAATTEAPGEYFFSDGIRLKKYRGMSLDAMDKHLSSONRYFSEA---DKIVAVGVGSA 446
DB 360 LAGTDESPEGLTYYGGRQYKSYRGMSLGAMOK--GSSDRYFOGTAODKL--VPEGIEGR 416
OY 447 VODKSIHKFVPLIAGIOHSCODIGAKSLTOVRAMYSGLKFEKRTSSAOVEGVHSL 506
DB 417 VVVGSIKRSVYHQLLGLGRSSMGYVGAKDIED----FQKRAFEVETITAGLAKESHVHDV 471

RESULT 13
O97FM8 PRELIMINARY: PRT: 485 AA.
AC O97FM8:
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
GN IMP dehydrogenase.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
  Gibson R., Lee H.M., Dubois J., Olu D., Hitti J., Wolf Y.I.,
  Tatusov R.L., Saathie F., Doucetle-Stamm L., Soucaille P., Daly M.J.,
```

RA Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
RT bacterium *Clostridium acetobutylicum*,"  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL: AE007768: AAK80647.1: -;  
DR InterPro: IPR000644: CBS\_domain.  
DR InterPro: IPR003009: FMN\_enzyme.  
DR InterPro: IPR001093: IMPDH/GMPrtase.  
DR Pfam: PF00571: CBS; 2.  
DR Pfam: PF00478: IMPDH\_C; 1.  
DR Pfam: PF01574: IMPDH\_N; 1.  
DR SMART: SM00116: CBS; 2.  
DR TIGRfams: TIGR01302: IMP\_dehydrog; 1.  
DR PROSITE: PS00487: IMP\_DH\_GMP\_RED; 1.  
KW Complete proteome.  
SQ SEQUENCE 485 AA: 52188 MW: 994DIC5347F7A267 CRC64;

Query Match 36.6%; Score 958; DB 16; Length 485;  
Best Local Similarity 44.2%; Pred. No. 1.4e-58;  
Matches 212; Conservative 86; Mismatches 164; Indels 18; Gaps 5;

QY 31 TYNDFLLPGYIDFTADVDLTLSALTKTKTLTKPLTVSSPMDVTVEAGMAIMALTGGTG 90  
DB 10 TFDDVLVPAVSAVLPRTSILRTKTRTELATPLVSAADVTVEARLAIMAQQGGIG 69  
QY 91 IHNCTPEFOANEVRKAKYKKEGFTTDPVYLSPKDRVDVEAKARHGFCGIPITDTGRM 150  
DB 70 IHNMTIAQASEVDPRKQENGVTNPISLSDNSVQALDLMKRYRLSGVPIIDN--- 126  
QY 151 GSRVGIISRDIDPLEEHDCFLLEIMTKREDLVVAARSTITKEANETLQRSKKGLP 210  
DB 127 AGRLLGITNRDIVF--ETDYSKRIEELMT--TENLTAQGTITIDEAKMLLKHKEKLP 183  
QY 211 IYNEDDELVAIIARTDLKKNRDYPLASKDAKKQLLGAIGTHEDDKYRLDLAQAQVD 270  
DB 184 LVDENFVLKGLTITIDIEIKRFPAKAKSQGRLLCGACAVGTAKMDKRYALVDASVY 243  
QY 271 VVLDSSQGSNIFQIMIKYIKDKYPNLOVIGSNVNTAAQAKNLIDAGVDALRVMGSGS 330  
DB 244 IYLDPAHGHSGQVLEAVKIKKAYPELQVTAQNVATAAIVHDLIEAGDCAVKGIGPSI 303  
QY 331 CIIQEVLAACGRPOATAVVYVEYARFGVPIADGIGQVNGHIAKALALGASTVMGSL 390  
DB 304 CTRRVAGIGVQPLTAVMDCEVANKYGPITADGIGIKYSGLIVALLAAGAAVMGSMF 363  
QY 391 AATTEAPGEYFFSDGIRLKKYRGMSLDAMDKHLSONRYFEADKIKYAQGVSGAVQDK 450  
DB 364 ACCEAPGETELIYQGRSYVYRGMGLAM--QCGSKDYEFGGKRLVPEGEVGPVK 421  
QY 451 GSIRKFPVYPLIAGIOHSCODIGAKSLTOVRAMYSSELKFEKRTSSAQVEGVSLSHYE 510  
DB 422 GSVEITVFOILLGIRSGMGLSRNMTL-----FEKATFVQVGSAGLRSHPHD 471

RESULT 14  
Q8XZG6 PRELIMINARY: PRT: 487 AA.

AC Q8XZG6;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Probable inosine-5'-monophosphate dehydrogenase oxidoreductase protein  
DE (EC 1.1.1.205).  
GN GUAB OR RSC1429 OR RS05263.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
OC Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GMT1000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

RA Ariat M., Billault A., Brotier P., Camus J.C., Cattolico L.,  
RA Chandiier M., Choise N., Claudel-Renard C., Cunne S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Siglier P., Thebaud P., Whalen M., Winker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"  
RL Nature 415:497-502(2002).  
DR EMBL: AL646064: CAD15131.1: -;  
DR InterPro: IPR000644: CBS\_domain.  
DR InterPro: IPR003009: FMN\_enzyme.  
DR InterPro: IPR001093: IMPDH/GMPrtase.  
DR Pfam: PF00571: CBS; 2.  
DR Pfam: PF00478: IMPDH\_C; 1.  
DR Pfam: PF01574: IMPDH\_N; 1.  
DR SMART: SM00116: CBS; 2.  
DR TIGRfams: TIGR01302: IMP\_dehydrog; 1.  
DR PROSITE: PS00487: IMP\_DH\_GMP\_RED; 1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 487 AA: 52065 MW: A4052A6119CF011 CRC64;

Query Match 35.9%; Score 941.5; DB 16; Length 487;  
Best Local Similarity 42.7%; Pred. No. 2e-57;  
Matches 206; Conservative 87; Mismatches 170; Indels 19; Gaps 6;

QY 30 LYNDFLIPGYIDFTADVDLTLSALTKTKTLTKPLTVSSPMDVTVEAGMAIMALTGGTG 89  
DB 8 LFDDVLVPAVSAVLPRTSILRTKTRTELATPLVSAADVTVEARLAIMAQQGGIG 67  
QY 90 FIHNCTPEFOANEVRKAKYKKEGFTTDPVYLSPKDRVDVEAKARHGFCGIPITDTGR 149  
DB 68 IYHKLKLPDEQAREVAKYKRFESGVLRDPIITGPPMKVARDVALSAQHIGSEFVLE--- 124  
QY 150 MCSRLVGIISRDIDPLEEHDCFLLEIMTKREDLVVAARSTITKEANETLQRSKKGL 209  
DB 125 -GNKYVGIITNDLRF--EELDAPVRAKMTGKELVTVREBASLDEKRLKMKRRLERV 181  
QY 210 PYNEDDELVAIIARTDLKKNRDYPLASKDAKKQLLGAIGTHEDDKYRLDLAQAQVD 269  
DB 182 LVYDGNFELRGLITKDIQATEHPLASKDERGSLRVGAAGVGPDPNDLRYDLVAKAVD 241  
QY 270 VVLDSSQGSNIFQIMIKYIKDKYPNLOVIGSNVNTAAQAKNLIDAGVDALRVMGSGS 329  
DB 242 VLVVPTAHGHSGQVLSRVKWLIDKTPQVOVIGGNATATAAALADHDGADGVKVGIGSS 301  
QY 330 CIIQEVLAACGRPOATAVVYVEYARFGVPIADGIGQVNGHIAKALALGASTVMGSL 389  
DB 302 IOTTRIVAGVGPQLSANVAEALAKNTGVPYADGAVRYSGDIAKALLAAGHVTMMGGM 361  
QY 390 LAATTEAPGEYFFSDGIRLKKYRGMSLDAMDKHLSONRYFE-----ADKIKYAGVS 444  
DB 362 FAGTEEAGEVFLYQGRSYKSYRGMGSVAMKD--GAADRYFQEDNFTANVDKL-VPEGIE 418  
QY 445 GAVQDGSIRKFPVYPLIAGIOHSCODIGAKSLTOVRAMYSSELKFEKRTSSAQVEGVH 504  
DB 419 GRVPYKGSVLPVHQLTGIRSMGYCGASTAE-----WBEKSFQVITAGMRESHVH 473

505 SL 506  
474 DV 475

RESULT 15  
Q9KGN8 PRELIMINARY: PRT: 485 AA.

AC Q9KGN8;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Inositol-monophosphate dehydrogenase (EC 1.1.1.205).  
GN GUAB OR BH0020.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.



```
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki H., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001507; BAB03739.1; -.
DR HSSP: P50099; 1ZFU.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPdh/GMPrtase.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PR01574; IMPDH_N; 1.
DR SMART: SM00116; CBS; 2.
DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 485 AA: 52429 MW: D0B64895CF8B60 CRC64:

Query Match 35.2%; Score 922; DB 16; Length 485;
Best local Similarity 40.2%; Pred. No. 4.4e-56;
Matches 196; Conservative 107; Mismatches 164; Indels 20; Gaps 6;

QY 28 DGLTYNDFLIPGIDFTADVDLTSAITKTKITPLVSSPMDVYERAGMATAMALTGG 87
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 9 EGLTFDDVLLPAPKSEVLPADVSVKTKITFTLNLNPIISAGMDVYERAKMAIAIREGG 68
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 88 IGFHNHCTPEFOANKREYKVKTEQGFITDPVVLSPKRVRYDFEAKARRHGFCIPITDT 147
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 69 LGIHNMSVEEQAEQVDYKRSSEGVITNPFLTPDQVFEDEHLMCKYRISGVPIVDE 128
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 148 GRMSRLVGIISSRDIDFLKEEEDCFLEIMTKREDLVVAPRSITLKEANEILORSKKG 207
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 129 DQ---KLVGILTRDLRFI--EDYSTLIDVMTK-ENLYTAPVGTILKEAEELQKHKE 182
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 208 KLPVNEDELVAIIARTDLKKNRDYPPLASKDAKKQLLCGAIGTHEIDKRYRLDLLAQA 267
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 183 KLPLVDESGTLKGLITIKDIEKYLEFPNSAKDSQGRILVGAAGVSAIDVRYAALVEAG 242
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 268 VDVVVLDSQSGNISFQINMKYIKDKYPNLOVIGVNVVTAQAQKNLIDAGVDALRVGMS 327
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 243 VDVIVIDTAHGHSGVLEKVKALREOYPLDTIAGNVATAEATRDLEIRGANVYKVGIGP 302
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 328 GSICITQEVLAGCRPOATAVYKVEYARFQVPIADGCIQNGHIAKALALGASTVMMG 387
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 303 GSICITRIVAGIGVPOITVAYDCANEARKHGVPITADGGIKYSGDLVKAALAGHAVMLG 362
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 388 SLIAATTEARGEYFFSDGIRLKYRGMGSLDAMDKHLSSONRYFSEADIKVAQVSGAV 447
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 363 SLIAGVSESPGEREIFQGRQFYRGMGSLGAWK--GSKDRYFOENNDKLVEPEGIEGRI 420
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 448 ODGSIHKFEVYPLIAGIOHSCODIGAKSLTOVAMMYSGELKFEKRTSBAOVGEG-CVHSL 506
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 421 PYKGPLHDTIHQLVGCIRAGMGYCGTKITDELK-----ENTQITRITGAGLRES 469
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 507 HSYEKRL 513
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 470 HPDHDVQI 476
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```

Search completed: February 13, 2003, 06:32:37  
Job time : 104 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2003, 01:37:31 : Search time 9 Seconds

(without alignment)  
546.637 Million cell updates/sec

Title: US-09-846-637C-4

Sequence: 1 MADVLISGRTSYRPPDGLTA.....SSAQVEGVMILHSYKRLP 514

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 11892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2595	99.1	514	1	IMD2_HUMAN
2	2559	97.7	514	1	P12268 homo sapien
3	2555	97.6	514	1	P12269 mesocricetu
4	2222	84.8	514	1	IMD2_MOUSE
5	2203	84.1	514	1	IMD1_HUMAN
6	1764.5	67.4	537	1	IMD1_MOUSE
7	1610.5	61.5	524	1	IMD4_DROME
8	1607.5	61.4	523	1	IMD4_YEAST
9	1605.5	61.3	521	1	IMH3_CANAL
10	1589.5	60.7	523	1	IMH3_YEAST
11	1473.5	56.3	524	1	IMD2_YEAST
12	1414	54.0	454	1	IMDH_SCHPO
13	1365.5	52.9	512	1	IMDH_PNECA
14	1371.5	52.4	514	1	IMDH_TRYBB
15	1155	44.1	502	1	IMDH_LEIDO
16	1106.5	42.2	503	1	IMH1_ARATH
17	1043.5	39.8	490	1	IMH1_ARATH
18	952	36.3	490	1	IMD1_YEAST
19	887.5	33.9	488	1	IMDH_AQUAE
20	857.5	32.7	488	1	IMDH_BACSU
21	849.5	32.4	529	1	IMDH_ACICA
22	839.5	32.0	529	1	IMDH_MYCTU
23	837	32.0	488	1	IMDH_MYCLE
24	836	31.9	488	1	IMDH_ECOLI
25	827.5	31.6	521	1	IMDH_HAEIN
26	811	31.0	481	1	IMDH_CHLVI
27	811	31.0	487	1	IMDH_HELPY
28	809	30.9	481	1	IMDH_PASMU
29	808	30.9	492	1	IMDH_HELPJ
30	772	29.5	498	1	IMDH_STRPY
31	770	29.4	496	1	IMDH_RHTRP
32	704.5	26.9	485	1	IMDH_METUA
33	689.5	26.3	485	1	IMDH_PYRPU
					IMDH_PYRAB

## ALIGNMENTS

RESULT 1	ID	IMD2_HUMAN	STANDARD	PRT:	514 AA.
AC	P12268				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-MAY-1991 (Rel. 18, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Inosine-5'-monophosphate dehydrogenase 2 (EC 1.1.1.205) (IMP dehydrogenase 2) (IMPDH-II) (IMPD 2).				
GN	IMPDH2 OR IMPD2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=89008491; PubMed=2902093;				
RA	Collart F.R., Huberman E.;				
RT	"Cloning and sequence analysis of the human and Chinese hamster				
RT	inosine-5'-monophosphate dehydrogenase cDNAs.";				
RL	J. Biol. Chem. 265:15769-15772(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Spleen;				
RX	MEDLINE=90203022; PubMed=1969416;				
RA	Natsuneda Y., Ohno S., Kawasaki H., Konno Y., Weber G., Suzuki K.;				
RT	"Two distinct cDNAs for human IMP dehydrogenase.";				
RL	J. Biol. Chem. 265:5292-5295(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Blood;				
RX	MEDLINE=95091778; PubMed=7999076;				
RA	Gleason D.A., Huberman E.;				
RT	"Cloning and sequence of the human type II IMP dehydrogenase gene.";				
RL	Biochem. Biophys. Res. Commun. 205:537-544(1994).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=95204479; PubMed=7896827;				
RA	Zimmerman A.G., Spyrcala J., Mitchell B.S.;				
RT	"Characterization of the human inosine-5'-monophosphate dehydrogenase				
RT	type II gene.";				
RL	J. Biol. Chem. 270:6808-6814(1995).				
RN	[5]				
RP	CHARACTERIZATION.				
RX	MEDLINE=95283610; PubMed=7763314;				
RA	Hager P.W., Collart F.R., Huberman E., Mitchell B.S.;				
RT	"Recombinant human inosine monophosphate dehydrogenase type I and				
RT	type II proteins. Purification and characterization of inhibitor				
RL	binding.";				
RN	Biochem. Pharmacol. 49:1323-1329(1995).				
RN	[6]				
RP	X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).				
RX	MEDLINE=99199217; PubMed=10097070;				
RA	Colby T.D., Vanderveen K., Strickler M.D., Markham G.D.,				
RA	Goldstein B.M.;				
RT	"Crystal structure of human type II inosine monophosphate				

34	684	26.1	486	1	IMDH_PYRHO	O56045 pyrococcus
35	618	23.6	404	1	IMDH_BORBU	P49058 borrelia bu
36	582.5	22.2	503	1	IMDH_TRIFO	P50097 tritrichomo
37	463	17.7	479	1	Y143_MYCTU	O50591 mycobacteri
38	437	16.7	478	1	Y143_MYCLE	O32912 mycobacteri
39	326	12.4	345	1	GUAC_HUMAN	P36959 homo sapien
40	323.5	12.4	356	1	GUAC_ASCSU	P27442 ascaris suu
41	317	12.1	345	1	GUAC_RAT	O92244 rattus norv
42	315.5	12.0	358	1	GUAC_CAEL	O16294 caenorhabdi
43	293	11.2	346	1	GUAC_ECOCI	P15344 escherichia
44	292	11.1	349	1	GUAC_BUCAI	P57300 buchnera ap
45	166	6.3	265	1	Y188_METUA	O57647 methanococc

```

RT dehydrogenase: implications for ligand binding and drug design.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3531-3536(1999).
CC -1- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
CC OF CELL GROWTH. IT MAY ALSO HAVE A ROLE IN THE DEVELOPMENT OF
CC MALIGNANCY AND THE GROWTH PROGRESSION OF SOME TUMORS.
CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
CC xanthosine 5'-phosphate + NADH.
CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- TISSUE SPECIFICITY: IMP TYPE I IS THE MAIN SPECIES IN NORMAL
CC LEUKOCYTES AND TYPE II PREDOMINATES OVER TYPE I IN THE TUMOR.
CC -1- PTM: The N-terminus is blocked.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, J04208; AAA6112.1; -.
CC EMBL, L33842; AAA67054.1; -.
CC EMBL, L39210; AAB70699.1; -.
CC PIR, A31997; A31997.
CC PIR, B35566; B35566.
CC PDB, 1B30; 12-APR-99.
CC Genew: HGNC:6053; IMPDH2.
CC MIM, 146691; -.
CC InterPro: IPR000644; CBS_domain.
CC InterPro: IPR003009; FMN_enzyme.
CC InterPro: IPR001093; IMPDH/GMPrtase.
CC Pfam: PF00478; IMPDH_C; 1.
CC Pfam: PF00571; CBS; 2.
CC Pfam: PF01574; IMPDH_N; 1.
CC SMART: SM00116; CBS; 2.
CC TIGRFAMS: TIGR01302; IMP_dehydrog; 1.
CC PROSITE: PS00487; IMP_DH_GMP_RED; 1.
CC Oxidoreductase; NAD: GMP biosynthesis; Purine biosynthesis;
CC Multigene family; Repeat; CBS domain; 3D-structure.
CC FT DOMAIN 112 167 CBS 1.
CC FT DOMAIN 177 232 CBS 2.
CC FT BINDING 331 331 IMP (POTENTIAL).
CC FT CONFLICT 190 191 AG -> RS (IN REF. 1).
CC SO SEQUENCE 514 AA; 55805 MW; 876BEA0EC1DBEE9 CRC64;
-----
Query Match 99.1%; Score 2595; DB 1; Length 514;
Best Local Similarity 99.2%; Pred. No. 3.6e-168;
Matches 510; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 301 GGNVTTAAQAKNLIDAGVDALRVGSGSICITQEVILACGRPOATAVYKYVEYARFGVP 360
DB 301 GGNVTTAAQAKNLIDAGVDALRVGSGSICITQEVILACGRPOATAVYKYVEYARFGVP 360
OY 361 VIADGGIONVGHIAKALAGASTVMGSLAATPAEPGFPSDGRILKRYGKMSLDAM 420
DB 361 VIADGGIONVGHIAKALAGASTVMGSLAATPAEPGFPSDGRILKRYGKMSLDAM 420
OY 421 DKHLSQNRFESEADKIRVAQVSGAVODKGSIHKEFVPLIANGIOHSCODIGAKSLTOVR 480
DB 421 DKHLSQNRFESEADKIRVAQVSGAVODKGSIHKEFVPLIANGIOHSCODIGAKSLTOVR 480
OY 481 AMMYSGELKFEKRTSSAQVEGVHSLHSYKRLF 514
DB 481 AMMYSGELKFEKRTSSAQVEGVHSLHSYKRLF 514
-----
RESULT 2
ID2_MESAU
ID IM2_MESAU STANDARD; PRT; 514 AA.
AC P12269;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase 2 (EC 1.1.1.205) (IMP
DE dehydrogenase 2) (IMPDH-II) (IMPD 2).
GN IMPDH2 OR IMPDH.
OS Mesocricetus auratus (golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 336-370.
RX MEDLINE=89008491; Pubmed=2902093;
RA Collart F.R., Huberman E.;
RT "Cloning and sequence analysis of the human and Chinese hamster
RT inosine-5'-monophosphate dehydrogenase cDNAs.";
RL J. Biol. Chem. 263:15769-15772(1988).
CC -1- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
CC OF CELL GROWTH. IT MAY ALSO HAVE A ROLE IN THE DEVELOPMENT OF
CC MALIGNANCY AND THE GROWTH PROGRESSION OF SOME TUMORS.
CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
CC xanthosine 5'-phosphate + NADH.
CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, J04209; AAA6993.1; -.
CC PIR, B31997; B31997.
CC HSSP: P12268; 1B30.
CC InterPro: IPR000644; CBS_domain.
CC InterPro: IPR003009; FMN_enzyme.
CC InterPro: IPR001093; IMPDH/GMPrtase.
CC Pfam: PF00478; IMPDH_C; 1.
CC Pfam: PF00571; CBS; 2.
CC Pfam: PF01574; IMPDH_N; 1.
CC SMART: SM00116; CBS; 2.
CC TIGRFAMS: TIGR01302; IMP_dehydrog; 1.
CC PROSITE: PS00487; IMP_DH_GMP_RED; 1.
CC Oxidoreductase; NAD: GMP biosynthesis; Purine biosynthesis;

```

KM Multigene family: Repeat: CBS domain.  
 FT DOMAIN 112 167 CBS 1.  
 FT DOMAIN 177 232 CBS 2.  
 FT BINDING 331 IMP (POTENTIAL).  
 SQ SEQUENCE 514 AA: 55890 MW: 5FA0138FA4E8A02 CRC64:

Query Match 97.7%: Score 2559; DB 1: Length 514;  
 Best local similarity 98.1%: Pred. No. 9, 7e-165;  
 Matches 504; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MADYLLISGTSYVPDDGLTAQQLFNCGDGLTYNDFLLPGYIDFTADVADLTSALTKKIT 60  
 DB 1 MADYLLISGTSYVPDDGLTAQQLFNCGDGLTYNDFLLPGYIDFTADVADLTSALTKKIT 60  
 QY 61 LKTPLVSSPMDDVTEAGMAIAMAALTGIGFIHNCTPEFOANVVRKYYEGGFITDPVY 120  
 DB 61 LKTPLVSSPMDDVTEAGMAIAMAALTGIGFIHNCTPEFOANVVRKYYEGGFITDPVY 120  
 QY 121 LSPKDRVDVFEAKARHGFCGIPITDTGMRGSRVGISSRDIDFLKEEHDFLEIMT 180  
 DB 121 LSPKDRVDVFEAKARHGFCGIPITDTGMRGSRVGISSRDIDFLKEEHDFLEIMT 180  
 QY 181 KREDLVVAPRSTILKANELLORSKKGLPIVNEDELVAIIARTDLKKNRDPPLASKDA 240  
 DB 181 KREDLVVAPRSTILKANELLORSKKGLPIVNEDELVAIIARTDLKKNRDPPLASKDA 240  
 QY 241 KQOLCGAAIGTHEDDKYRLDLAAGVDVYVLDSSQGSNIFQINMIYIKDKYPNLOVY 300  
 DB 241 KQOLCGAAIGTHEDDKYRLDLAAGVDVYVLDSSQGSNIFQINMIYIKDKYPNLOVY 300  
 QY 301 GGNVVTAAQAKNLDAGVALRVGSGSICITQEVLAGCRPATAV KYEYARRFQVY 360  
 DB 301 GGNVVTAAQAKNLDAGVALRVGSGSICITQEVLAGCRPATAV KYEYARRFQVY 360  
 QY 361 VIADGCIQNVGHIAKALALGASTVMGSLAATTEAPGEFFSDGIR KKYRGMGSLDAM 420  
 DB 361 VIADGCIQNVGHIAKALALGASTVMGSLAATTEAPGEFFSDGIR KKYRGMGSLDAM 420  
 QY 421 DKHLSSQNTYFSEADKTKAAGVSGAVQDKGSIHKFVPTLNGTOS ODIGKSLTVOR 480  
 DB 421 DKHLSSQNTYFSEADKTKAAGVSGAVQDKGSIHKFVPTLNGTOS ODIGKSLTVOR 480  
 QY 481 AMYSGELKFEKRTSSAOGVGVHSLHSEYKRLF 514  
 DB 481 AMYSGELKFEKRTSSAOGVGVHSLHSEYKRLF 514

RESULT 3  
 ID IMD2\_MOUSE STANDARD: PRT: 514 AA.  
 AC P24547: 061734;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 RT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Inosine-5'-monophosphate dehydrogenase 2 (EC 1.1.1.205) (IMP  
 GN dehydrogenase 2) (IMPDH-1) (IMPD 2).  
 OS IMPDH2 OR IMPDH.  
 OC Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN RP SEQUENCE FROM N.A., AND VARIANTS MYCOPHENOLIC ACID RESISTANT.  
 RA MEDLINE=9115361; PubMed=1671845;  
 RT Tiedeman A.A., Smith J.M.;  
 RL "Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase";  
 [2]  
 Gene 97:289-293(1991).  
 RP SEQUENCE FROM N.A., AND VARIANTS MYCOPHENOLIC ACID RESISTANT.  
 RC TISSUE=Brain;  
 RA MEDLINE=9415391; PubMed=7906545;  
 RL Lightfoot T., Snyder F.F.;  
 RT "Gene amplification and dual point mutations of mouse IMP

RT dehydrogenase associated with cellular resistance to mycophenolic  
 acid.";  
 RL Biochim. Biophys. Acta 1217:156-162(1994).  
 CC -1- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS  
 CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION  
 CC OF CELL GROWTH. IT MAY ALSO HAVE A ROLE IN THE DEVELOPMENT OF  
 CC MALIGNANCY AND THE GROWTH PROGRESSION OF SOME TUMORS.  
 CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =  
 CC xanthosine 5'-phosphate + NADH.  
 CC -1- PATHWAY: FIRST REACTION UNIT TO GMP BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
 CC GMP REDUCTASE.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M33934; AAA39311.1;  
 CC EMBL: M98333; AAA20181.1;  
 CC PIR: J70565; J70565.  
 CC DR HSSP: P12268; 1B30.  
 CC DR PMMA-2DPAGE: P24547;  
 CC DR MGD: MGI:109367; Impdh2.  
 CC InterPro: IPR000644; CBS domain.  
 CC InterPro: IPR003009; FMN-enzyme.  
 CC InterPro: IPR001093; IMPDH/GMPylase.  
 CC Pfam: PF00478; IMPDH\_C; 1.  
 CC Pfam: PF00571; CBS; 2.  
 CC Pfam: PF01574; IMPDH\_N; 1.  
 CC SMART: SM00116; CBS; 2.  
 CC TIGRfam: TIGR01302; IMP dehydrog; 1.  
 CC PROSITE: PS00487; IMP DH\_GMP\_RED; 1.  
 CC Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;  
 CC Multigene family: Repeat: CBS domain.  
 KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;  
 KM Multigene family: Repeat: CBS domain.  
 FT DOMAIN 112 167 CBS 1.  
 FT BINDING 177 232 CBS 2.  
 FT BINDING 331 331 IMP (POTENTIAL).  
 FT VARIANT 333 333 T -> I (IN MYCOPHENOLIC ACID RESISTANT  
 FT VARIANT 351 351 S -> Y (IN MYCOPHENOLIC ACID RESISTANT  
 FT VARIANT 351 351 CELLS).  
 FT CONFLICT 483 483 T -> M (IN REF. 2).  
 FT SEQUENCE 514 AA: 55785 MW: D5B66A5C5EBCC421 CRC64:

Query Match 97.6%: Score 2555; DB 1: Length 514;  
 Best local similarity 97.9%: Pred. No. 1, 8e-165;  
 Matches 503; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MADYLLISGTSYVPDDGLTAQQLFNCGDGLTYNDFLLPGYIDFTADVADLTSALTKKIT 60  
 DB 1 MADYLLISGTSYVPDDGLTAQQLFNCGDGLTYNDFLLPGYIDFTADVADLTSALTKKIT 60  
 QY 61 LKTPLVSSPMDDVTEAGMAIAMAALTGIGFIHNCTPEFOANVVRKYYEGGFITDPVY 120  
 DB 61 LKTPLVSSPMDDVTEAGMAIAMAALTGIGFIHNCTPEFOANVVRKYYEGGFITDPVY 120  
 QY 121 LSPKDRVDVFEAKARHGFCGIPITDTGMRGSRVGISSRDIDFLKEEHDFLEIMT 180  
 DB 121 LSPKDRVDVFEAKARHGFCGIPITDTGMRGSRVGISSRDIDFLKEEHDFLEIMT 180  
 QY 181 KREDLVVAPRSTILKANELLORSKKGLPIVNEDELVAIIARTDLKKNRDPPLASKDA 240  
 DB 181 KREDLVVAPRSTILKANELLORSKKGLPIVNEDELVAIIARTDLKKNRDPPLASKDA 240  
 QY 241 KQOLCGAAIGTHEDDKYRLDLAAGVDVYVLDSSQGSNIFQINMIYIKDKYPNLOVY 300  
 DB 241 KQOLCGAAIGTHEDDKYRLDLAAGVDVYVLDSSQGSNIFQINMIYIKDKYPNLOVY 300

```

OY 301 GGNVYTAQAQKMLIDAGVDAIRVGMGSGSICITIOEVLACGRPOATVYKVEYARREGVF 360
DB 301 GGNVYTAQAQKMLIDAGVDAIRVGMGSGSICITIOEVLACGRPOATVYKVEYARREGVF 360
OY 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEYFFSDGIRLKKYRGMGLDM 420
DB 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEYFFSDGIRLKKYRGMGLDM 420
OY 421 DKHLSSQNRFFSEADKIKYAAQGVSGAVODKGIHFKFVPLIAGIOHSCODIGAKSLTOYR 480
DB 421 DKHLSSQNRFFSEADKIKYAAQGVSGAVODKGIHFKFVPLIAGIOHSCODIGAKSLTOYR 480
OY 481 AMYSGELKFEKRTSSAQVEGCVHSLHSEKRLF 514
DB 481 AMYSGELKFEKRTSSAQVEGCVHSLHSEKRLF 514

RESULT 4
IMD1_HUMAN STANDARD; PRT; 514 AA.
AC P20839;
DT 01-FEB-1991 (rel. 17, Created)
DT 01-FEB-1991 (rel. 17, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase 1 (EC 1.1.1.205) (IMP
DE dehydrogenase 1) (IMPDH-1) (IMPD 1).
GN IMPDH1 OR IMPD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RX MEDLINE=90203022; PubMed=1969416;
RA Natsumeda Y., Ohno S., Kawasaki H., Konno Y., Weber G., Suzuki K.;
RT "Two distinct cDNAs for human IMP dehydrogenase.";
RL J. Biol. Chem. 265:5292-5295(1990).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95283610; PubMed=7763314;
RA Hager P.W., Collart F.R., Huberman E., Mitchell B.S.;
RT "Recombinant human Inosine monophosphate dehydrogenase type I and
RT type II proteins. Purification and characterization of inhibitor
RT binding.";
RL Biochem. Pharmacol. 49:1323-1329(1995).
CC -!- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
CC OF CELL GROWTH. IT MAY ALSO HAVE A ROLE IN THE DEVELOPMENT OF
CC MALIGNANCY AND THE GROWTH PROGRESSION OF SOME TUMORS.
CC -!- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
CC xanthosine 5'-phosphate + NADH.
CC -!- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: IMP TYPE I IS THE MAIN SPECIES IN NORMAL
CC LEUKOCYTES AND TYPE II PREDOMINATES OVER TYPE I IN THE TUMOR.
CC -!- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05272; AAA36114.1; ALT_SEQ.
DR PIR: A35566; A35566.
DR HSP: P12268; I830.
DR GeneW: HGNC:6052; IMPDH1.

```

```

DR MW: 146690; -
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPDH/GMPrtase.
DR Pfam: PF00478; IMPDH_C; 1.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF01574; IMPDH_N; 1.
DR SMART: SM00116; CBS; 2.
DR TIGRfams: TIGR01302; IMP dehydrog; 1.
DR PROSITE: PS00487; IMP DH GMP RED; 1.
KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;
KW Multigene family; Repeat; CBS domain.
FT DOMAIN 112 167 CBS 1.
FT DOMAIN 177 232 CBS 2.
FT BINDING 331 331 IMP (POTENTIAL).
SQ SEQUENCE 514 AA; 55449 MW; 751303F8E9E21061 CRC64;

Query Match 84.8%; Score 2222; DB 1; Length 514;
Best Local Similarity 83.1%; Pred. No. 5,9e-143;
Matches 427; Conservative 39; Mismatches 48; Indels 0; Gaps 0;

OY 1 MADYLISSGTSYVPDDGLTAQOLFNGDGLTYNDPLILPGYIDFTADQVLDLSALTKIT 60
DB 1 MADYLISSGTSYVPDDGLTAQOLFASADLTYNDFLIFGFIIDFADDEVLTLSALTKIT 60
OY 61 LKTPLYSSPMDTVTEAGMAIAMLTGIGIFHHNCTPEFOANEVRKKYKKEGFTIDPV 120
DB 61 LKTPLYSSPMDTVTEADMAIAMLMGIGIFHHNCTPEFOANEVRKKYKKEGFTIDPV 120
OY 121 LSPKADRVYFPAKARHGCIGIPITDTGRMGRSLVGISSRIDLPKKEEHDFLEIWT 180
DB 121 LSPKADRVYFPAKARHGCIGIPITDTGRMGRSLVGISSRIDLPKKEEHDFLEIWT 180
OY 121 LSPKADRVYFPAKARHGCIGIPITDTGRMGRSLVGISSRIDLPKKEEHDFLEIWT 180
DB 121 LSPKADRVYFPAKARHGCIGIPITDTGRMGRSLVGISSRIDLPKKEEHDFLEIWT 180
OY 181 KREDLVAPRSTLKPANELLORSKKGLPIYNEDELVAIARTDKKRNRPPLASKDA 240
DB 181 KREDLVAPRSTLKPANELLORSKKGLPIYNEDELVAIARTDKKRNRPPLASKDA 240
OY 181 PRELVVAPAGVTLKANELLORSKKGLPIYNDCELVAIARTDKKRNRPPLASKDS 240
DB 181 PRELVVAPAGVTLKANELLORSKKGLPIYNDCELVAIARTDKKRNRPPLASKDS 240
OY 241 KQQLLGAALIGTHEDDKRYRLDLAAGVDVYVLDSSQSGNSFOINMKYIKDKYPLQYT 300
DB 241 KQQLLGAALIGTHEDDKRYRLDLAAGVDVYVLDSSQSGNSFOINMKYIKDKYPLQYT 300
OY 241 KQQLLGAALIGTHEDDKRYRLDLAAGVDVYVLDSSQSGNSFOINMKYIKDKYPLQYT 300
DB 241 KQQLLGAALIGTHEDDKRYRLDLAAGVDVYVLDSSQSGNSFOINMKYIKDKYPLQYT 300
OY 301 GGNVYTAQAQKMLIDAGVDAIRVGMGSGSICITIOEVLACGRPOATVYKVEYARREGVF 360
DB 301 GGNVYTAQAQKMLIDAGVDAIRVGMGSGSICITIOEVLACGRPOATVYKVEYARREGVF 360
OY 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEYFFSDGIRLKKYRGMGLDM 420
DB 361 VIADGGIOWNGHIAKALALGASTVMMGSLAATTEAPGEYFFSDGIRLKKYRGMGLDM 420
OY 421 DKHLSSQNRFFSEADKIKYAAQGVSGAVODKGIHFKFVPLIAGIOHSCODIGAKSLTOYR 480
DB 421 DKHLSSQNRFFSEADKIKYAAQGVSGAVODKGIHFKFVPLIAGIOHSCODIGAKSLTOYR 480
OY 481 AMYSGELKFEKRTSSAQVEGCVHSLHSEKRLF 514
DB 481 AMYSGELKFEKRTSSAQVEGCVHSLHSEKRLF 514

RESULT 5
IMD1_MOUSE STANDARD; PRT; 514 AA.
AC P50096;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase 1 (EC 1.1.1.205) (IMP
DE dehydrogenase 1) (IMPDH-1) (IMPD 1).
GN IMPDH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.  
RC TISSUE=leukemia;  
RA Dayson J.S., Mitchell B.S.;  
RL Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: IMP is the RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS  
CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION  
CC OF CELL GROWTH. IT MAY ALSO HAVE A ROLE IN THE DEVELOPMENT OF  
CC MALIGANCY ACTIVITY. THE GROWTH PROGRESSION OF SOME TUMORS.  
CC -!- CATALYTIC ACTIVITY: inosine 5'-phosphate + NAD(+) + H(2)O =  
CC xanthosine 5'-phosphate + NADH.  
CC -!- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
CC GMP REDUCTASE.  
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U00978; AAA18285.1; -  
CC DR HSSP: P12268; 1B30.  
DR MGD: MG1:96567; Impdh1.  
DR InterPro: IPR000644; CBS\_domain.  
DR InterPro: IPR001093; FMN\_enzyme.  
DR pfam: PF00478; IMPDH\_C; 1.  
DR pfam: PF00571; CBS; 2.  
DR pfam: PF01574; IMPDH\_N; 1.  
DR SMART: SM00116; CBS; 2.  
DR TIGRfams: TIGR01302; IMP\_dehydrog; 1.  
DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; 1.  
KW Oxidoreductase; NAD; IMP; DH GMP; RED; 1.  
KW Multigene family; Repeat; CBS domain.  
FT DOMAIN 112 167 CBS 1.  
FT 177 232 CBS 2.  
FT BINDING 331 IMP (POTENTIAL).  
SQ SEQUENCE 514 AA; 55294 MW; B38A2EB0ECC4CE CRC64;  
  
Query Match 84.1%; Score 2203; DB 1; Length 514;  
Best Local Similarity 82.5%; Pred. No. 1; Le-141;  
Matches 424; Conservative 39; Mismatches 51; Gaps 0; Gaps 0;  
  
OY 1 MADYLSGGTSTVPPDDGLTAQQLFNCGGDLTYNDFLLPGYIDFTADYDLTSALTKKIT 60  
DB 1 MADYLSGGTGYVPEDGLTAHFLFANADGLTYNDFLLPGYIDFTADYDLTSALTKKIT 60  
OY 61 LKTPLVSSMDPTVTEAGMALAMALTTGGIGFIHNCPTPEFOANEVRKAEYEGFTDPVV 120  
DB 61 LKTPLVSSMDPTVTEAGMALAMALTTGGIGFIHNCPTPEFOANEVRKAEYEGFTDPVV 120  
OY 121 LSPKORVDVFEAKARHFGCGIPITDTGMRGSLVGISSRDIIDFLKEEHDCFILEIMT 180  
DB 121 LSPKORVDVFEAKARHFGCGIPITDTGMRGSLVGISSRDIIDFLKEEHDCFILEIMT 180  
OY 181 KREDLVAPRSTLKEANFILLORSKKGLPIVNEDELVAIIARIDLKNNRPYPLASKXA 240  
DB 181 KREDLVAPRSTLKEANFILLORSKKGLPIVNEDELVAIIARIDLKNNRPYPLASKXA 240  
OY 241 KROLLCGAATGTHEDDKYRLDLAAGVDDVYLLDSOGNSIFOINNIYTKDKYPLNOYI 300  
DB 241 KROLLCGAATGTHEDDKYRLDLAAGVDDVYLLDSOGNSIFOINNIYTKDKYPLNOYI 300  
OY 301 GGNVYTAAGAKNLLDAGVALRVGMGSGSICITIOEVLACGRQATAV KVEYARFEGVP 360  
DB 301 GGNVYTAAGAKNLLDAGVALRVGMGSGSICITIOEVLACGRQATAV KVEYARFEGVP 360  
OY 361 VIADGCIQVGHIAKALALGASTVMGSLLAATTEAPGEFFSDGIRKKYRGMGLDAM 420  
DB 361 VIADGCIQVGHIAKALALGASTVMGSLLAATTEAPGEFFSDGIRKKYRGMGLDAM 420

DB 361 VIADGCIQVGHIAKALALGASTVMGSLLAATTEAPGEFFSDGIRKKYRGMGLDAM 420  
OY 421 DKHLSSQNRPFSEADKIKVAGCVGAVYDKSTHFRVPLYLAGIQHSCODIGAKSLTOVR 480  
DB 421 EKSSSQKRFYFSGEDVKVIAQGVSGSIDPKSGIYOKFRVPLYLAGIQHSCODIGAKSLTOVR 480  
OY 481 AMYSGELKEFKRTSSAOYEGVGHSHSYEKRLF 514  
DB 481 SMYSGELKEFKRTSSAOYEGVGHSHSYEKRLF 514  
  
RESULT 6  
IDMD DROME STANDARD: PRT: 537 AA.  
AC 007152: Q26455: Q9W2R8:  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP  
DE dehydrogenase) (IMPDH) (IMPD) (Raspberry protein).  
GN RAS OR CG1799.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
(1)  
RP SEQUENCE FROM N.A.  
RA Nash D., Hu S.;  
RT "Drosophila inosine monophosphate dehydrogenase is encoded at the  
RT raspberry locus.";  
RL (in) Abstracts of the 35th meeting of the Canadian Federation of  
RL Biological Societies, pp.72-72, Victoria (1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94259281; PubMed=7911114;  
RA Nash D., Hu S., Leonard N.J., Tjong S.Y., Phillips D.;  
RT "The raspberry locus of Drosophila melanogaster includes an inosine  
RT monophosphate dehydrogenase like coding sequence.";  
RL Genome 37:333-344(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=canlon-S;  
RX MEDLINE=94114565; PubMed=7904480;  
RA Sifri C.D., Wilson K., Smolik S., Forte M., Ullman B.;  
RT "Cloning and sequence analysis of a Drosophila melanogaster cDNA  
RT encoding IMP dehydrogenase.";  
RL Biochim. Biophys. Acta 1217:103-106(1994).  
RN [4]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RX MEDLINE=96069715; PubMed=7476879;  
RA Slee R., Bownes M.;  
RT "The raspberry locus encodes Drosophila inosine monophosphate  
RT dehydrogenase.";  
RL Mol. Gen. Genet. 248:755-766(1995).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie J.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Chang M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.A., Bouck J., Brokslein P., Brothier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jajili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirsks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -i- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS  
 CC OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION  
 CC OF CELL GROWTH.  
 CC -i- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =  
 CC xanthosine 5'-phosphate + NADH.  
 CC -i- PATHWAY: FIRST REACTION UNIOUE TO GMP BIOSYNTHESIS.  
 CC -i- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
 CC GMP REDUCTASE.  
 CC -i- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL: L14847; AAA21831.1; -  
 DR EMBL: L22608; AAA16839.1; -  
 DR EMBL: S80430; AAB35628.1; -  
 DR EMBL: AE003451; AAF46622.1; -  
 DR HSSP: P12268; 1B30.  
 DR FLYBASE: FBgn0003204; ras.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR InterPro: IPR003009; FMN\_enzyme.  
 DR InterPro: IPR001093; IMPDH/GMPrtase.  
 DR Pfam: PF00478; IMPDH\_C; 1.  
 DR Pfam: PF00571; CBS; 2.  
 DR Pfam: PF01574; IMPDH\_N; 1.  
 DR SMART: SM00116; CBS; 2.  
 DR TIGRFAMS: TIGR01302; IMP\_dehydrog; 1.  
 DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; 1.  
 KM Oxidoreductases; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;  
 KM CBS domain.  
 FT DOMAIN 134 189 CBS 1.  
 FT 198 251 CBS 2.  
 FT BINDING 350 350 IMP (BY SIMILARITY).  
 FT CONFLICT 38 38 D -> V (IN REF. 4).  
 FT CONFLICT 53 53 T -> P (IN REF. 4).  
 FT CONFLICT 99 102 EMAI -> RCH (IN REF. 4).  
 FT CONFLICT 184 184 D -> A (IN REF. 4).  
 FT CONFLICT 194 194 V -> S (IN REF. 4).  
 FT CONFLICT 216 217 AN -> EH (IN REF. 4).  
 FT CONFLICT 226 229 GLPL -> ATA (IN REF. 4).  
 FT CONFLICT 244 244 T -> A (IN REF. 4).  
 FT CONFLICT 261 262 KO -> TR (IN REF. 4).  
 FT CONFLICT 266 266 VG -> CP (IN REF. 4).  
 FT CONFLICT 277 278 AR -> GCRA (IN REF. 4).  
 CC -----

FT CONFLICT 284 284 A -> R (IN REF. 4).  
 FT CONFLICT 301 301 Y -> I (IN REF. 4).  
 FT CONFLICT 387 388 OS -> HA (IN REF. 4).  
 SQ SEQUENCE 537 AA: 57829 MW: A5EA6A1AEAA64EBD CWC64;  
 Query Match 67.4%; Score 1764.5; DB 1; Length 537;  
 Best Local Similarity 68.0%; Pred. No. 5,2e-112;  
 Matches 342; Conservative 72; Mismatches 82; Indels 7; Gaps 3;  
 16 DGLTQQLPFGCDGTYNDFLPGYIDFTADQVDFSLTAKITKLPVSSPDYTE 75  
 |||:::|||||  
 38 DGLSKELQNEGLTYNDFLPGYIDFTAEVDLSPLTSLTRADLVSSPDYTE 97  
 |||:::|||||  
 76 AGMAIMALTGIGIFHNCTPEFOANEVKKYKKEOGFITDPVVLSPKRDVFEAKA 135  
 |||:::|||||  
 98 SEMAIALMALCGSIGIHNCTPEYQALEVHKYKKGFMRDPSVSPINTGVLEARR 157  
 |||:::|||||  
 136 RIGFGIPITPDGKMSKLVGISSRDIDFLKEEHDCFLIEIKRKREDLVVAPSTLK 195  
 ::|||:::|||||  
 158 KNGFGYPYTGKGLGKLLGWTSRIDF-RENQPEVLADIIM--TELVPNGINLP 214  
 |||:::|||||  
 196 ENETIQRKSKKLPVNEDEDLVAIARTDLKKNRDYPLASKDAKKOLLGAAIGTHED 255  
 |||:::|||||  
 215 TNAILEKSKKLPVNOAGELVAMARTDLKARSTPMASKDSKKOLLGAAIGTISE 274  
 |||:::|||||  
 256 DRYRLDLAQAQGVVYVLDSSQNSIFQIMTKYIKDKXPNIQVIGNVYTAQAQKMLID 315  
 |||:::|||||  
 275 DKARLALLAVANGVDVILIDSSQNSVYQVEMIKYIKETYPDELQVIGNVYTAQAQKMLID 334  
 |||:::|||||  
 316 AGVDALRVGMSGSTICIIDVYLACGPRQATVAYKYEVARRGCVYIADGIQNVGHTAK 375  
 |||:::|||||  
 335 AGVDGLRVGMSGSTICIDQVYMACCPQATVAYQVSTARQFVPIADGGIQTSGHTYK 394  
 |||:::|||||  
 376 ATALGATVMGSLAATPEARPEYFSDGIRLKKYRGSGISDAMD---KHLSSQNRFF 431  
 |||:::|||||  
 395 ATALGASAVMGSLLAGTSEAPGEITFSDGVRLKTRGSGLEAERGDAAKAAASRYTH 454  
 |||:::|||||  
 432 SEADIKYAGVSGAVODKSTHKRPVYLIAIQHSCODIGAKSLTOVRAAMYSGELKE 491  
 :|||:::|||||  
 455 NEMDKKVAQVSGSIVDGSVLTARYPLECGLQSCODIGANSINKLADMIYNGQLRM 514  
 |||:::|||||  
 492 KRTSSAOVGGVSHLSHYEKRLP 514  
 |||:::|||||  
 515 KRTHSAQLEGVNHGLFSYEKRLE 537  
 |||:::|||||  
 RESULT 7  
 IMD4\_YEAST  
 ID IMD4\_YEAST STANDARD: PRT: 524 AA.  
 AC P50094;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable inosine-5'-monophosphate dehydrogenase IMD4 (EC 1.1.1.205)  
 DE (IMP dehydrogenase) (IMPDH) (IMPD).  
 GN IMD4 OR YM01056C OR YM9958.06C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxId=4932;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;  
 RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.  
 CC -i- CATALYTIC ACTIVITY: to the EMBL/Genbank/DBJ databases.  
 CC xanthosine 5'-phosphate + NADH.  
 CC -i- PATHWAY: FIRST REACTION UNIOUE TO GMP BIOSYNTHESIS.  
 CC -i- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
 CC GMP REDUCTASE.  
 CC -i- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 246729; CAA6719.1; -  
DR HSSP: P12268; 1B30.  
DR SGP: S0004520; IMD4.  
DR InterPro: IPR000644; CBS\_domain.  
DR InterPro: IPR003009; FMN\_enzyme.  
DR InterPro: IPR001093; IMPDH/GMPrtase.  
DR Pfam: PF00478; IMPDH\_C; 1.  
DR Pfam: PF00571; CBS; 2.  
DR Pfam: PF01574; IMPDH\_N; 1.  
DR SMART: SM00116; CBS; 2.  
DR TIGRfams: TIGR01302; IMP\_dehydrog; 1.  
DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; 1.  
KM Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;  
KM Multigene family; Repeat; CBS domain.  
FT DOMAIN 120 175 CBS 1.  
FT DOMAIN 183 235 CBS 2.  
FT BINDING 336 336 IMP (POTENTIAL).  
SQ SEQUENCE 524 AA: 56394 MW; A73DIEFEF8AEAD9 CRC64

Query Match 61.5%; Score 1610.5; DB 1; Length 524;  
Best Local Similarity 61.8%; Pred. No. 1.3e-101;  
Matches 316; Conservative 82; Mismatches 102; Indels 11; Gaps 5;

QY 11 SYVPDGLTAQOLF--CGDGLTYNDLFLPGYIDFTAOVDLTSLTKITLKTPLVSS 68  
DB 17 TTSQDGLVQLDMLSTTRGTLTYNDFLPGLVNFPSSAVSLQTLTKITLTPFVSS 76  
QY 69 PMDTVEAGMATAMLTSGIGFIHNHCTPEFOANEVRYKKYKKEOGFTLPVLSPKDRVR 128  
DB 77 PMDTVEADMAIYMLLGLGIGFIHNHCTPEFOASMYKKYKMFENGINPIVISPTTVG 136  
QY 129 DYFEKARKGFCGIPITDGRMGSLVGIISRRDIFLKEEHDCLEIRIKREDLVYA 188  
DB 137 EYKVMKRGKGFSGFPVTEGKCPGKLVGLVTSRDIOFL--EDDSLVSIVMTKNP--VTG 192  
QY 189 PSRITLKEANELLORSKKGLPIVNEDELVAIARTDLKKNRDPYAKKDA-KKOLLGC 247  
DB 193 IKGITLKEENELIKQTKKKLLIVDNGNLVMSLRADLMKNONPLAKSATTKOLLGC 252  
QY 248 AALGHEDDKYLLDLAAGVYVYVLDSSGNSITQINMIKTIKDYPLQVYIGNVATA 307  
DB 253 AALGTEADKERLLRLLVEAGLDVILDSQNSVFOQNMKIKETFPLEIITAGNVATR 312  
QY 308 AQAANLIDAGVALRVGMGSGSICITIOEVLACGPOATVYVYVEYARFEGVPIADGCI 367  
DB 313 EGAANLIDAGAGGLIRLIGMGSSICITIOEVMACGROGTAVVNCOPANFEGVPMADGCV 372  
QY 368 ONVGHIAKALAGASTVWMSGLLAATTEAPGEYFSSQITRLKKYKMGGLDM---DKH 423  
DB 373 ONIGHITKALAGSSVWGMGLAGTTESPGEYFYKDKRLAYKMGGLIDAMOKTGKMG 432  
QY 424 LSSORRYFEADKIKYACVSGAVODKGIHKFVPLIINGIHSODITAKSLTQVRAM 483  
DB 433 NASTSRYSSESIVLYACVSGAVYDKGISIKKFIPLTYGLHSCODICESLITLKENV 492  
QY 484 VSGELKEFKRTSSAAGEGVSHLSHYERKFL 514  
DB 493 ONGEVRFERTASAOLEGVHNLHSYERKLY 523

RESULT 8  
IMD3\_YEAST  
ID IMD3\_YEAST STANDARD: PRT: 523 AA.  
AC P50095;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable inosine-5'-monophosphate dehydrogenase IMD3 (EC 1.1.1.205)  
DE (IMP dehydrogenase) (IMPDH) (IMPD).  
GN IMD3 OR YLR432W OR U9753.4.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972.  
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J.,  
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,  
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,  
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,  
RA Ritten L., Riles L., Tatch A., Trevisan E., Vignati D.,  
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =  
CC xanthosine 5'-phosphate + NADH.  
CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.  
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
CC GMP REDUCTASE.  
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U21094; AAB67516.1; -  
DR HSSP: P12268; 1B30.  
DR SGP: S0004424; IMD3.  
DR InterPro: IPR000644; CBS\_domain.  
DR InterPro: IPR003009; FMN\_enzyme.  
DR InterPro: IPR001093; IMPDH/GMPrtase.  
DR Pfam: PF00478; IMPDH\_C; 1.  
DR Pfam: PF00571; CBS; 2.  
DR Pfam: PF01574; IMPDH\_N; 1.  
DR SMART: SM00116; CBS; 2.  
DR TIGRfams: TIGR01302; IMP\_dehydrog; 1.  
DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; 1.  
KM Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;  
KM Multigene family; Repeat; CBS domain.  
FT DOMAIN 119 174 CBS 1.  
FT DOMAIN 182 234 CBS 2.  
FT BINDING 335 335 IMP (POTENTIAL).  
SQ SEQUENCE 523 AA: 56584 MW; AOC84C2257AAE6 CRC64

Query Match 61.4%; Score 1607.5; DB 1; Length 523;  
Best Local Similarity 62.2%; Pred. No. 2.1e-101;  
Matches 314; Conservative 81; Mismatches 99; Indels 11; Gaps 5;

QY 16 DGLTAQOLF--CGDGLTYNDLFLPGYIDFTADVDLTSLTKITLKTPLVSSPMDTV 73  
DB 21 DGLSVQELMDSKTRGTLTYNDFLPGLVDFPSSEVSLQTKLRTNLTITVPVSSPMDTV 80  
QY 74 TEAGMATAMLTSGIGFIHNHCTPEFOANEVRYKKYKKEOGFTTDPVLSPKDRVDFEA 133  
DB 81 TESEMAIFPALGIGIGFIHNHCTPEFOADMVRVRYKENGFINNPVISPTTVGEAKSM 140  
QY 134 KARHGCGLPIIDTGRMGSLVGIISRRDIFLKEEHDCLEIETKREDLVVAPRST 193  
DB 141 KERFGSGFPVTEGKRNKLMGIVTSRDIOY--EDNSLVQDVMTKNP--VTGAQGIT 196  
QY 194 LKEANELLORSKKGLPIVNEDELVAIARTDLKKNRDPYASKDA-KKOLLGAAIGT 252  
DB 197 LSEGNELIKKIKKKLLIVDNGNLVMSLRADLMKNONVPLASKSATTKOLLGAAIGT 256



RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaudin M.,  
 RT \*Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII.\*  
 RL Science 265:2077-2082(1994).  
 CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =  
 CC xanthosine 5'-phosphate + NADH.  
 CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.  
 CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
 CC GMP REDUCTASE.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U00029; AAB69728.1; -  
 CC PIR: S48997; S48997.  
 CC HSSP: P12268; 1B30.  
 CC SGD: S0001259; IMD2.  
 CC InterPro: IPR000644; CBS\_domain.  
 CC InterPro: IPR003009; FMN\_enzyme.  
 CC InterPro: IPR001093; IMPDH/GMPase.  
 CC Pfam: PF00478; IMPDH\_C; 1.  
 CC Pfam: PF00571; CBS\_2.  
 CC DR Pfam: PF01574; IMPDH\_N; 1.  
 CC SMART: SM00116; CBS\_2.  
 CC DR TIGRFAMs: TIGR01302; IMP\_dehydrog; 1.  
 CC DR PROSITE: PS00487; IMP\_DH\_GMP\_RED; 1.  
 CC KW oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;  
 KM Multigene family; Repeat; CBS domain.  
 FT DOMAIN 119 174 CBS 1.  
 FT DOMAIN 162 234 CBS 2.  
 FT BINDING 335 335 IMP (POTENTIAL).  
 SQ SEQUENCE 523 AA; 56530 MW; 7CA3EC11238906B9 CRC64

Query Match 60.7%; Score 1589.5; DB 1; Length 523;  
 Best local similarity 61.2%; Pred. No. 34e-100;  
 Matches 309; Conservative 84; Mismatches 101; Indels 11; Gaps 5;

QY 16 DGLTTAQLFN--CGDGLTYNDLILPGYIDFTADQVDTLSATLTKITLTPLVSSPMQTV 73  
 DB 21 DGLSVLELNDKIRGLTYNDLILPGYIDFTADQVDTLSATLTKITLTPLVSSPMQTV 80  
 QY 74 TEAGMAIAMLTGIGIFIRHNCIPFEQANFVRKVKYEGGFTIDPVYLTPKDRVRPFEEA 133  
 DB 81 TESEMATFALLGLGIGIFIRHNCIPFEQADMRVKYKENGFINNPVLPPTTVGSAKSM 140  
 QY 134 KARHGGCGIPITDTGMRGRLVGIISSRIDLKEEHDCLFEIMTKEDLVVAPRSTT 193  
 DB 141 KERYGAGPVPYITDCKRNKLVGITSRDIOFV--EDNSLIVQDVMTKIP--VTGAQGIT 196  
 QY 194 LKEANEILDRSKKGLPIYNEDDELVAIIARTDLKKNRYPPLASKRKKKOLLGCAIGT 252  
 DB 197 LSEGNLILKKIKRLLVYDEKGNLVSMLSRDLMKNQYPLASKRANKOLLGCAISITG 256  
 QY 253 HEDDKYRLDLAAGYDVVLDSSQGNSTFQINIMTKYIKDKYPNLOVITGNNVTAQAQAN 312  
 DB 257 MOADERELRLVKAGIDVILDSQGNSTFELNMLKWKVESFGLGVITGNNVTRQOAN 316  
 QY 313 LIDAGDALRVGMGSGSITIOEVLAGCRPOATVAVKYVYARRGCVPTADGGINVGH 372  
 DB 317 LIAAGADGIRIGMGTSICITQEVMAKGRPOGTAVYNVEFANQGVPTADGGINVGH 376  
 QY 373 TAKALALAGSTVWMSLAATTEAPGEYFSGCIRLTKTRGMSLDAMKHLSSON 428  
 DB 377 ITRKALALGSSVWMSGLAGTTESPEYFYODGKRLKAYRGMGSDIAMKKTGKNASTS 436  
 QY 429 RFSEADKIKVAGVSGAVQDKGSIHKEVPYLIAGTGHSCODGAKSLQVRAVMYSGEL 488

DB 437 RFSEDSVLLVAGVSGAVVQDKGSIKRPILYNGLQHSQDIDGCRSLTLKNVGRKV 496  
 QY 489 KFEKRTSSAÖVEGGVSHLSYERKL 513  
 DB 497 RFEFRTASAOEGGVHNLHSYERKL 521

RESULT 11  
 IDMDH\_SCHPO  
 ID 014344; Q9P3X8; STANDARD; PRT; 524 AA.  
 AC 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP  
 DE dehydrogenase) (IMPDH) (IMPD).  
 GN GDAL OR SPIC2F12.14C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972.  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Wolckaert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welfjens I., Vansireels E., Rieger M., Schaefer M., Wellert-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleuvre V., Motlier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 RN [2]  
 RC SEQUENCE OF 378-524 FROM N.A.  
 RC STRAIN=972.  
 RA Karier S., Topal Sarykaya A., Temizkan G.;  
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Essential enzyme in the de novo biosynthesis of GMP,  
 CC catalysing the NAD-dependent oxidation of IMP to XMP.  
 CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =  
 CC xanthosine 5'-phosphate + NADH.  
 CC -1- PATHWAY: First reaction unique to GMP biosynthesis.  
 CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO GMP  
 CC REDUCTASE.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

```
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: Z97211; CAB10161.1; -.
CC      DR EMBL: AJ293460; CAB97003.1; -.
CC      DR HSSP: P12268; 1B30.
CC      DR InterPro: IPR000644; CBS_domain.
CC      DR InterPro: IPR003009; PMN_enzyme.
CC      DR InterPro: IPR001093; IMPdh/GMPPrase.
CC      DR Pfam: PF00571; CBS; 2.
CC      DR Pfam: PF00478; IMPDH_C; 1.
CC      DR Pfam: PF01574; IMPDH_N; 1.
CC      DR SMART: SM00116; CBS; 2.
CC      DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
CC      KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis;
CC      Multigene family; Repeat; CBS domain.
CC      FT DOMAIN 119 175 CBS 1.
CC      FT BINDING 184 237 CBS 2.
CC      FT BINDING 337 337 IMP (POTENTIAL).
CC      FT CONFLICT 508 509 IR -> TRK (in REF. 2).
CC      FT SEQUENCE 524 AA; 57026 MW; E6C822C22E74674F CRC64;
SQ
Query Match
Best Local Similarity 56.3%; Score 1473.5; DB 1; Length 524;
Matches 296; Conservative 77; Mismatches 128; Indels 9; Gaps 4;
QY 12 YPPDGLTAAQQLF--NCGDGLVYNDFLLPGYIDFTADQVLTSLTKTKITPLVSSP 69
DB 17 YEKKDLSTIDLLIRHNFGGLTFENDFLLPGYIDFPVNNVSLFETRLSRNLVLTCPMSSP 76
QY 70 MDVTVEAGMALAMALGGTGFHNNCTPREQANEVKRYKKEGGFTTDPVYLSPKRDV 129
DB 77 MDVTVEQAMATYVALLGGIGVHNNCTPREQAMVKKRYKKEGGFTTDPVFSPOHTVGD 136
QY 130 VFEAKARHGFCGIPITDTGRMSRLVGISSRDIDFLKEEHDCELEIMTKREDLVAP 189
DB 137 VLKIKETKSGSIPITENCKLRKGLVTSRVOYFHKDINTP--VTEVNTPREBELITTA 194
QY 190 RSTTLEANEILLQSKKKKLPVNEDEDLVAIIARTDLKKRDYPLASKDA-KKOLLG 248
DB 195 EGSLEERANEMLRKSKKGLPVVDKDNVALLSTDLMKNLHFPLASKTSDFKQLMVA 254
QY 249 AITGHEHDDKYRDLDAQAGVYVLDSSQNSIFQINMKTITDKYPNQVIGCNVYTA 308
DB 255 AITGRDDDRIRLLAEAGDAVYVLDSSQNSCEQLEMTIKTKYTPKIDVAGNVYTR 314
QY 309 QAKNLIDAGVDALRVGMSSGICITIOEVLACGRPQATAVYKVEYARRFGVPYIADG 368
DB 315 QTSLLAAGADGLRVGMSSGACITIOEVACGRPQATATAVAFASQFCIGVYIADG 374
QY 369 NGCHIKALALGASTVWMSLLAATTEAPGEYFFSDGIRLKKYRGMSLDAMD---KHL 424
DB 375 NGCHMYKSLSLGATVWMSGLLAGTSPGEYVYREGORRYKRYGMSILAAMEGTGVN 434
QY 425 SSQNRFESEADRIKVAQVSGAVQDKSIHKFPVYLLAGIOHSCODIGKSLQVAMMY 484
DB 435 ASIGRIFSENDAVRAQVSGVLYVDKSLRLRPYLYTGDLHALDYGKSLDELAEVAD 494
QY 485 SGELKEFKRTSSAQVSGVHSLSYEKRLF 514
DB 495 KHEVREFELRSSAAIRBGDIQGFATYERKLY 524
```

```
OC      Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC      Pneumocystis.
OC      NCBI_TaxID=4754;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97135213; PubMed=8980752;
RA      O'gara M.J., Lee C.H., Weinberg G.A., Nott J.M., Queener S.F.;
RT      "IMP dehydrogenase from Pneumocystis carinii as a potential drug
RT      target.";
RL      Antimicrob. Agents Chemother. 41:40-48(1997).
CC      -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
CC      xanthosine 5'-phosphate + NADH.
CC      -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC      -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC      GMP REDUCTASE.
CC      -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U42442; AAA97462.1; -.
CC      DR HSSP: P12268; 1B30.
CC      DR InterPro: IPR000644; CBS_domain.
CC      DR InterPro: IPR003009; PMN_enzyme.
CC      DR InterPro: IPR001093; IMPdh/GMPPrase.
CC      DR Pfam: PF00571; IMPDH_C; 1.
CC      DR Pfam: PF00571; CBS; 2.
CC      DR Pfam: PF01574; IMPDH_N; 1.
CC      DR SMART: SM00116; CBS; 2.
CC      DR TIGRfams: TIGR01302; IMP_dehydrog; 1.
CC      DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
CC      KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;
CC      CBS domain.
CC      FT DOMAIN 47 102 CBS 1.
CC      FT BINDING 109 163 CBS 2.
CC      FT BINDING 263 263 IMP (BY SIMILARITY).
CC      FT SEQUENCE 454 AA; 49522 MW; 5E4B64426A730947 CRC64;
SQ
Query Match
Best Local Similarity 54.0%; Score 1414; DB 1; Length 454;
Matches 283; Conservative 67; Mismatches 94; Indels 14; Gaps 5;
QY 66 VSSPMDTVTEAGMALAMALTGIGTFHNNCTPREQANEVKRYKKEGGFTTDPVYLSPK 125
DB 1 MSSPMDTVTESMAINALALGGIGVHNNCTLEQREVRKRYKKEGGFTTSPVLSLNH 60
QY 126 RVADVEAKARHGFCGIPITDTGRMSRLVGISSRDIDFLKEEHDCELEIMTKREDL 185
DB 61 RRVADYRIKEELGFSIPITDTGQNGKLGLVTSRDIOFHNND--SLSEYITK--DL 116
QY 186 VVAPRSITLKEANEILLQSKKKKLPVNEDEDLVAIIARTDLKKRDYPLASK-DARKQL 244
DB 117 VTGSEGRIFLEFANEILLQSKKKKLPVDEKGNLTFALLSRDLNHLPLASKLPDSKQL 176
QY 245 LCGAATGHEHDDKYRDLDAQAGVYVLDSSQNSIFQINMKTITDKYPNQVIGCNVY 304
DB 177 ICAQAAGTRPDRIRIRKLHVEAGLDIVLDSSQNSIYQINNIKKNKKEFPMLEVIAGNV 236
QY 305 VTAQAQKNLIDAGVDALRVGMSSGICITIOEVLACGRPQATAVYKVEYARRFGVPYIAD 364
DB 237 VTRQQAANLISGADALRVGMSSGICITIOELMAVGRQATAVYVASEPASKFQVPTIAD 296
QY 365 GGIIQNVGHITAKALGASTVWMSLLAATTEAPGEYFFSDGIRLKKYRGMSLDAMDRL 424
DB 297 GGIIENGHITKALGASVWMSGLLAGTSPGQYYRRDGRQLKSYRGMSIDAME-HL 355
QY 425 SSQN-----RYFSEADRIKVAQVSGAVQDKSIHKFPVYLLAGIOHSCODIGKASL 476
```

```

Db 356 SCNNKGDNAASSRYFGEADRTIRVAQGVSSVIDKSLHVVYPLRTGIDHSLDIDIGVNL 415
OY 477 TOVRAMMYSGELKFEKRTSSAQVEGCVHSLHSEKRLF 514
Db 416 TELRKOVKEKNIRFEPRTVASQLEGNVHGLDSYOKKLM 453

RESULT 13
IMDH_TRYBB STANDARD: PRT: 512 AA.
ID IMDH_TRYBB
AC P50058;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP
OS dehydrogenase) (IMPDH) (IMPD).
OC Trypanosoma brucei brucei.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BATRO 164 / ISNat 1.7;
RX MEDLINE=95050714; PubMed=7961861;
RA Wilson K., Berens R.L., Sifri C.D., Ullman B.;
RT "Amplification of the inosinate dehydrogenase gene in Trypanosoma
brucei gambiense due to an increase in chromosome copy number.";
RL J. Biol. Chem. 269:28979-28987(1994).
CC -1- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
OF CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
xanthosine 5'-phosphate + NADH.
CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
GMP REDUCTASE.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL: M97794; AAB6420.1; -
DR HSSP: P12268; 1B30.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR001093; IMPDH/GMPrtase.
DR Pfam: PF00478; IMPDH_C. 1.
DR Pfam: PF00571; CBS_2.
DR Pfam: PF01574; IMPDH_N. 1.
DR SMART: SM00116; CBS; 2.
DR TIGRfams: TIGR01302; IMP_dehydrog. 1.
DR PROSITE: PS00487; IMP_DH_GMP_RED; 1.
KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;
KW CBS domain.
FT DOMAIN 108 163 CBS 1.
FT DOMAIN 170 226 CBS 2.
FT BINDING 325 325 IMP (POTENTIAL).
SQ SEQUENCE 512 AA; 55708 MW; 1A86C46AE6445045 CRC64

Query Match 52.9%; Score 1385.5; DB 1; Length 512;
Best Local Similarity 54.1%; Pred. No. 2e-86;
Matches 270; Conservative 91; Mismatches 135; Indels 3; Gaps 2;

OY 16 DGLTAQQLNCGGLTYNDFLLPGYIDFTADVDLTSLTKKTLTKTLLVSSPMDTVTE 75
DB 13 DGLTAQQLNCGGLTYNDFLLPGYIDFTADVDLTSLTKKTLTKTLLVSSPMDTVTE 71
OY 76 AGMAIMALTGIGTFHINCTPEFOANERVKKYKVEOGFTDPVLESPDRVADVEAKA 135

```

```

Db 72 SSMARMAALMGIGVLIHNNCTVEQARMRVSVKLYNRNGFTMKPKSVSPDVPSTINIKS 131
OY 136 RHGFCGIPITDTGRMGSRVLGIISRDIDFLKEEHDCELEIMTYREDLVAPRSITLK 195
Db 132 EKSGIGLTYIGKYGKGLIGVCTKTDIDFVKDASAP--VSQYMTRENNMTVERYP1KLE 189
OY 196 EANEILLORSKKGKPLTVNEDEDLVAIARFDLKNNDYPLASDKAKKLLGSAIGTHED 255
Db 190 EAMVDLNRSHGTLPVLNDEDEVVCLSRDAVARARYPNSSIDRNGHLCAAAATSTREA 249
OY 256 DKYRLDLAAGVADVVYVLDSSQGNISFQINMIKYIKDKYPNLQVGNVTAQAANLID 315
Db 250 DKGRVALSEAGIDVLYVLDSSQGNITTYQVSEFIRWVKTPYHLEVVAGNVVTOQAANLID 309
OY 316 AGVDALRYMGSSGSLTIOVYLVACGRPQRAVYKVEYARRPVPYIADGIONVGHIAK 375
Db 310 AGADSLRIGMGSSGSLTIOVYLVACGRPQATAYKVARVYASRGVPCVADGLRNVDVCK 369
OY 376 ALALGASTVMGSLAATTEAPEGEFFSDGIRLKKYRGMSLDAMDHLSSQNRFEAD 435
Db 370 ALAVGANVAMLGSMIAGTSETPEYFFKQMRKGYRGMGSDAMLGREGSKRYLSENE 429
OY 436 KIVVAGVSAVDKSGIRKFFVYLLAGIOHSCODIGASLTOVRAMMYSGELKFEKRTS 495
Db 430 TLQVAGVADAVYLDKGSVYLLAYIHKGLOQSAODIGEVSFDAIRKRYEGVGLFNRRTL 489
OY 496 SAQVEGVSHLSYERKLF 514
Db 490 TAOSEGAHVSHHYERKLF 508

RESULT 14
IMDH_TRYBB STANDARD: PRT: 514 AA.
ID IMDH_TRYBB
AC P21620;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP
OS dehydrogenase) (IMPDH) (IMPD).
OS Leishmania donovani.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5661;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91107664; PubMed=1671039;
RA Wilson K.E., Collart F.R., Huberman E., Stringer J.R., Ullman B.;
RT "Amplification and molecular cloning of the IMP dehydrogenase gene of
Leishmania donovani."
RL J. Biol. Chem. 266:1665-1671(1991).
CC -1- FUNCTION: IMP IS THE RATE LIMITING ENZYME IN THE DE NOVO SYNTHESIS
OF GUANINE NUCLEOTIDES AND THEREFORE IS INVOLVED IN THE REGULATION
OF CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
xanthosine 5'-phosphate + NADH.
CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
GMP REDUCTASE.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL: M55667; AAA29253.1; -
DR PIR: A38668; A38668.
DR HSSP: P12268; 1B30.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR003009; FMN_enzyme.

```





THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2003, 04:36:17 ; Search time 33 seconds

(without alignments)  
932.323 Million cell updates/sec

Title: US-09-846-637C-4

Perfect score: 2619  
Sequence: 1 MADYLISGTSYVDDGLTA.....SSAOVGGVFSLHSEYKRLF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2595	99.1	514	A31997	IMP dehydrogenase
2	2559	97.7	514	B31997	IMP dehydrogenase
3	2555	97.6	514	JT0565	IMP dehydrogenase
4	2222	84.8	514	A35566	IMP dehydrogenase
5	1764.5	67.4	537	S41064	IMP dehydrogenase
6	1635.5	62.4	537	S59508	IMP dehydrogenase
7	1610.5	61.4	524	S50890	IMP dehydrogenase
8	1607.5	61.4	523	S59402	IMP dehydrogenase
9	1589.5	60.7	523	S48997	IMP dehydrogenase
10	1473.5	56.3	524	T40127	IMP dehydrogenase
11	1385.5	52.9	514	A55407	IMP dehydrogenase
12	1363.5	52.1	514	A38668	IMP dehydrogenase
13	1324.5	50.6	509	T32709	IMP dehydrogenase
14	1155	44.1	492	F86298	IMP dehydrogenase
15	1106.5	42.2	503	JC4999	IMP dehydrogenase
16	1043.5	39.8	403	S53477	IMP dehydrogenase
17	966.5	36.9	485	G81308	IMP dehydrogenase
18	958	36.6	485	D97232	IMP dehydrogenase
19	952	36.3	490	H70473	IMP dehydrogenase
20	922	35.2	485	D83652	IMP dehydrogenase
21	914	34.9	487	F81906	IMP dehydrogenase
22	911	34.8	487	H81109	IMP dehydrogenase
23	910	34.7	485	B82558	IMP dehydrogenase
24	904.5	34.5	482	C72264	IMP dehydrogenase
25	886.5	33.8	488	DEBSMP	IMP dehydrogenase
26	885	33.8	509	JC7305	IMP dehydrogenase
27	875	33.4	489	H83173	IMP dehydrogenase
28	860	32.8	488	G89805	IMP dehydrogenase
29	858.5	32.8	499	AC3621	IMP dehydrogenase

30	857.5	32.7	488	1	S23226	IMP dehydrogenase
31	855.5	32.7	500	1	F75342	IMP dehydrogenase
32	849.5	32.4	529	1	H70736	IMP dehydrogenase
33	846.5	32.3	488	2	AE1419	IMP dehydrogenase
34	846.5	32.3	488	2	AF1794	IMP dehydrogenase
35	846	32.3	489	2	C82282	IMP dehydrogenase
36	839.5	32.1	529	1	S72823	IMP dehydrogenase
37	837	32.0	488	1	DECIP	IMP dehydrogenase
38	837	32.0	488	2	B91050	IMP dehydrogenase
39	837	32.0	488	2	G85894	IMP dehydrogenase
40	836	31.9	488	2	H64055	IMP dehydrogenase
41	832.5	31.8	493	2	D86652	IMP dehydrogenase
42	827.5	31.6	521	1	T17196	IMP dehydrogenase
43	827	31.6	487	2	AC0349	IMP dehydrogenase
44	819.5	31.3	487	2	H87449	IMP dehydrogenase
45	818	31.2	490	2	AE0820	IMP dehydrogenase

## ALIGNMENTS

RESULT 1  
A31997  
IMP dehydrogenase (EC 1.1.1.205) II - human  
N:Alternate names: inosine 5'-monophosphate dehydrogenase (IMPDH) type II  
C:Species: Homo sapiens (man)  
C:Date: 21-May-1990 #sequence-revision 22-May-1998 #text-change 16-Jul-1999  
C:Accession: I52303; I54184; A92676; B35566; A31997; A94550  
R:Gleeson, D.A.; Huberman, E.  
Biochem. Biophys. Res. Commun. 205, 537-544, 1994  
A:Title: Cloning and sequence of the human type II IMP dehydrogenase gene.  
A:Reference number: I52303; MIMD:95091778; PMID:7959076  
A:Accession: I52303  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-514 <GLE1>  
A:Cross-references: GB:I33842; NID:602457; PIDN:AAA67054.1; PID:602458  
R:Gleeson, D.A.; Collart, F.R.; Varkony, T.; Drabkin, H.; Huberman, E.  
Genomics 16, 274-277, 1993  
A:Title: Chromosomal localization and structure of the human type II IMP dehydrogenase  
A:Reference number: I54184; MIMD:9322398; PMID:8098009  
A:Accession: I54184  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 461-514 <GLE2>  
A:Cross-references: GB:I08114; NID:9292239; PIDN:AAA36113.1; PID:9252240  
R:Collart, F.R.; Huberman, E.  
J. Biol. Chem. 263, 15769-15772, 1988  
A:Title: Cloning and sequence analysis of the human and Chinese hamster inosine-5'-mo  
A:Reference number: A92676; MIMD:89008491; PMID:2902093  
A:Accession: A92676  
A:Molecule type: mRNA  
A:Residues: 1-189, 'RS', 192-514 <COL>  
A:Cross-references: GB:J04208; NID:9186391; PIDN:AAA36112.1; PID:9307066  
R:Note: Submitted to the Protein Sequence Database, November 1989  
R:Natsumeda, Y.; Ohno, S.; Kawasaki, H.; Konno, Y.; Weber, G.; Suzuki, K.  
J. Biol. Chem. 265, 5292-5295, 1990  
A:Title: Two distinct cDNAs for human IMP dehydrogenase.  
A:Reference number: A35566; MIMD:90203022; PMID:1969416  
A:Accession: B35566  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-514 <MPT>  
C:Comment: mRNA for IMP dehydrogenase I (see PIR:A35566) predominated in normal leuko  
C:Genetics:  
A:Gene: GDB:IMPDH2  
A:Cross-references: GDB:128086; OMIM:146691  
A:Map position: 3p24.2-3p21.2  
A:Enrons: 33/2; 49/3; 83/3; 108/3; 177/3; 207/1; 274/1; 304/1; 336/1; 432/2; 480/2;  
C:Function:  
A:Description: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecu  
A:Pathway: purine nucleotide biosynthesis  
C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homo

C:Keywords: GMP biosynthesis; NAD: oxidoreductase; purine nucleotide biosynthesis  
 F:30-96/Domain: IMP dehydrogenase amino-terminal homology <IDHN>  
 F:117-168/Domain: CBS homology <CBS1>  
 F:184-232/Domain: CBS homology <CBS2>  
 F:233-476/Domain: IMP dehydrogenase catalytic homology <IDHC>  
 F:331/Active site: Cys #status Predicted

Query Match 99.1%; Score 2595; DB 1; Length 514;  
 Best Local Similarity 99.2%; Pred. No. 3.2e-160;  
 Matches 510; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MADYLLISGGSYVPDDGLTAQQLFNCGDGLTYNDPILPGYIDFTADQVDTLSALTKKIT 60
    |||||
DB 1 MADYLLISGGSYVPDDGLTAQQLFNCGDGLTYNDPILPGYIDFTADQVDTLSALTKKIT 60
    |||||
QY 61 LKPLVSSPMDDYVTEAGMAIAMAALTGGIGFIHNCIPEPQANVRRVKYEGGFTIDPVV 120
    |||||
DB 61 LKPLVSSPMDDYVTEAGMAIAMAALTGGIGFIHNCIPEPQANVRRVKYEGGFTIDPVV 120
    |||||
QY 121 LSPKDRVRFVEFAKARHCGIPITDTGRMGSRLVGISSRDIDFLKEEHDCLFEIIMT 180
    |||||
DB 121 LSPKDRVRFVEFAKARHCGIPITDTGRMGSRLVGISSRDIDFLKEEHDCLFEIIMT 180
    |||||
QY 181 KREDLVVAPRSITLKEANEILORSKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
    |||||
DB 181 KREDLVVAPRSITLKEANEILORSKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
    |||||
QY 241 KQOLCGAALGTHEDKRYRLDLAAGVDPVVLDSQGSIFQINMKYIKDKYPMLOYI 300
    |||||
DB 241 KQOLCGAALGTHEDKRYRLDLAAGVDPVVLDSQGSIFQINMKYIKDKYPMLOYI 300
    |||||
QY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICIIOEVLACGRPATAVYKVEYARFEGVP 360
    |||||
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICIIOEVLACGRPATAVYKVEYARFEGVP 360
    |||||
QY 361 VIADGSIQVNGHTAKALALGASTVMMGSLAATTEAPGEYFSDGIRLKKYRMGSLDAM 420
    |||||
DB 361 VIADGSIQVNGHTAKALALGASTVMMGSLAATTEAPGEYFSDGIRLKKYRMGSLDAM 420
    |||||
QY 421 DKHLSSONRFFSEADKIKVAQVSGAVODKGSIHKFPVPLIAGIOHSCODIGAKSLTOVR 480
    |||||
DB 421 DKHLSSONRFFSEADKIKVAQVSGAVODKGSIHKFPVPLIAGIOHSCODIGAKSLTOVR 480
    |||||
QY 481 AMMYSGELKFEKRTSSAOYEGVHSLHSYERKLF 514
    |||||
DB 481 AMMYSGELKFEKRTSSAOYEGVHSLHSYERKLF 514
    |||||

```

## RESULT 2

B31997  
 IMP dehydrogenase (EC 1.1.1.205) - Chinese hamster  
 N:Alternate names: IMP:NAD+ oxidoreductase; IMPDH; Inosine-5'-monophosphate dehydrogenase  
 C:Species: Cricetus griseus (Chinese hamster)  
 C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 05-May-2000  
 C:Accession: B31997  
 R:Collart, F.R.; Huberman, E.  
 J. Biol. Chem. 263, 15769-15772, 1988  
 A:Title: Cloning and sequence analysis of the human and Chinese hamster inosine-5'-monophosphate dehydrogenase  
 A:Reference number: A92676; MUID:89008491; PMID:2902093  
 A:Accession: B31997  
 A:Molecule type: mRNA  
 A:Residues: 1-514 <COLD>  
 A:Cross-references: GB:J04209; NID:g191119; PIDN:AAA36993.1; PID:g304517  
 C:Function:  
 A:Description: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecule of water to form inosine 5'-phosphate and NADH  
 A:Pathway: purine nucleotide biosynthesis  
 C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homology  
 C:Keywords: duplication; GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosyn  
 F:30-96/Domain: IMP dehydrogenase amino-terminal homology <IDHN>  
 F:117-168/Domain: CBS homology <CBS1>  
 F:184-232/Domain: CBS homology <CBS2>  
 F:233-476/Domain: IMP dehydrogenase catalytic homology <IDHC>  
 F:331/Active site: Cys #status Predicted

Query Match 97.7%; Score 2559; DB 1; Length 514;  
 Best Local Similarity 98.1%; Pred. No. 6.9e-158;  
 Matches 504; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 MADYLLISGGSYVPDDGLTAQQLFNCGDGLTYNDPILPGYIDFTADQVDTLSALTKKIT 60
    |||||
DB 1 MADYLLISGGSYVPDDGLTAQQLFNCGDGLTYNDPILPGYIDFTADQVDTLSALTKKIT 60
    |||||
QY 61 LKPLVSSPMDDYVTEAGMAIAMAALTGGIGFIHNCIPEPQANVRRVKYEGGFTIDPVV 120
    |||||
DB 61 LKPLVSSPMDDYVTEAGMAIAMAALTGGIGFIHNCIPEPQANVRRVKYEGGFTIDPVV 120
    |||||
QY 121 LSPKDRVRFVEFAKARHCGIPITDTGRMGSRLVGISSRDIDFLKEEHDCLFEIIMT 180
    |||||
DB 121 LSPKDRVRFVEFAKARHCGIPITDTGRMGSRLVGISSRDIDFLKEEHDCLFEIIMT 180
    |||||
QY 181 KREDLVVAPRSITLKEANEILORSKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
    |||||
DB 181 KREDLVVAPRSITLKEANEILORSKKGKLPYNEDELVAIARTDLKKNRDYPLASKDA 240
    |||||
QY 241 KQOLCGAALGTHEDKRYRLDLAAGVDPVVLDSQGSIFQINMKYIKDKYPMLOYI 300
    |||||
DB 241 KQOLCGAALGTHEDKRYRLDLAAGVDPVVLDSQGSIFQINMKYIKDKYPMLOYI 300
    |||||
QY 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICIIOEVLACGRPATAVYKVEYARFEGVP 360
    |||||
DB 301 GGNVVTAAQAKNLIDAGVDALRVGMGSGSICIIOEVLACGRPATAVYKVEYARFEGVP 360
    |||||
QY 361 VIADGSIQVNGHTAKALALGASTVMMGSLAATTEAPGEYFSDGIRLKKYRMGSLDAM 420
    |||||
DB 361 VIADGSIQVNGHTAKALALGASTVMMGSLAATTEAPGEYFSDGIRLKKYRMGSLDAM 420
    |||||
QY 421 DKHLSSONRFFSEADKIKVAQVSGAVODKGSIHKFPVPLIAGIOHSCODIGAKSLTOVR 480
    |||||
DB 421 DKHLSSONRFFSEADKIKVAQVSGAVODKGSIHKFPVPLIAGIOHSCODIGAKSLTOVR 480
    |||||
QY 481 AMMYSGELKFEKRTSSAOYEGVHSLHSYERKLF 514
    |||||
DB 481 AMMYSGELKFEKRTSSAOYEGVHSLHSYERKLF 514
    |||||

```

## RESULT 3

JT0565  
 IMP dehydrogenase (EC 1.1.1.205) - mouse  
 N:Alternate names: IMP:NAD+ oxidoreductase; IMPDH; Inosine-5'-monophosphate dehydrogenase  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 05-May-2000  
 C:Accession: JT0565; A34375; S42724  
 R:Riedeman, A.A.; Smith, J.M.  
 Gene 97, 289-293, 1991  
 A:Title: Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase.  
 A:Reference number: JT0565; MUID:91153661; PMID:1671845  
 A:Accession: JT0565  
 A:Molecule type: mRNA  
 A:Residues: 1-514 <TIB>  
 A:Cross-references: GB:M33934; NID:9198393; PIDN:AAA39311.1; PID:g309413  
 R:Hodges, S.D.; Fung, E.; McKay, D.J.; Renaux, B.S.; Snyder, F.F.  
 J. Biol. Chem. 264, 18137-18141, 1989  
 A:Title: Increased activity, amount, and altered kinetic properties of IMP dehydrogenase  
 A:Reference number: A34375; MUID:90036890; PMID:2572589  
 A:Accession: A34375  
 A:Molecule type: protein  
 A:Residues: 125-134; 187-194; 289-290; 439-449; 456-457, 'L', 459-464, 'S', 466; 475-478 <NOD>  
 R:Lightfoot, T.; Snyder, F.F.  
 Biochim. Biophys. Acta 1217, 156-162, 1994  
 A:Title: Gene amplification and dual point mutations of mouse IMP dehydrogenase assoc  
 A:Reference number: S42724; MUID:94153991; PMID:7906545  
 A:Accession: S42724  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-332, 'T', 334-350, 'Y', 352-482, 'M', 484-514 <LIG>  
 A:Cross-references: EMBL:M98333; NID:g425157; PIDN:AAA20181.1; PID:g425158









A:Residues: 1-512 <MIL>  
 A:Cross-references: GB:M9794; NID:9162135; PIDN:AAB6420.  
 A:Note: authors translated the codon CCG for residue 76 as  
 C:Genetics:  
 A:Gene: Impdh  
 C:Function:  
 A:Description: catalyzes the oxidation of inosine 5'-phosph  
 A:Pathway: purine nucleotide biosynthesis  
 C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydr  
 C:Keywords: GMP biosynthesis; NAD; oxidoreductase; purine  
 F:26-92/Domain: IMP dehydrogenase amino-terminal homology  
 F:177-228/Domain: CBS homology <CBS>  
 F:227-470/Domain: IMP dehydrogenase catalytic homology <ID  
 F:325/Active site: Cys #status predicted

Query Match 52.9%; Score 1385.5; DB 1; Length 512;  
 Best Local Similarity 54.1%; Pred. No. 5e-82;  
 Matches 270; Conservative 91; Mismatches 135; Indels 3; Gaps 2;

```

OY 16 DGLTAQOLNCGDGLTYNDFLIPGYIDPTAOVDLTSLTKRKLTKLT
DB 13 DCGTAEELFR-ODGLSFNDFILPGFIDPDSSKVNVSQGFKNILHL
OY 76 AGMAIAMLTGIGIFHNCTPEFOANEVRYKRYEKGFTDPVYLSPKDRVDFEAKA 135
DB 72 SSMARAMALMGIGIYHNCTVEEQARWRSYKLYRNGFTMKPKSVSP
OY 136 RHGFCGIPITDTRGMSRLVGISSNDIDFLKEEHDCLFEIMTKRE
DB 132 EKGISGILVTENGDPHKGILGIVCTKIDIDYVANK--DTPVSAVMTRREKMYERAPITOLE 191
OY 196 EANEILORSKKGKLPVNEDEDLVALIARTDLKKNRDYPLASKDAKKOLLGCAAGTCHED 255
DB 190 EAMVDLNRSGRYGLPVNENDEVNLCSRDVARRADYPSHTLDSGRILCAAASTREPE 251
OY 256 DKRYRLDLAQAQVDVYVLDSSQGNSTFQINMTIKYKDKPNIQVIGNVYTAQAOKNLID 315
DB 252 DKRYRLDLAQAQVDVYVLDSSQGNSTFQINMTIKYKDKPNIQVIGNVYTAQAOKNLID 311
OY 316 AGVDALRVMGSGSICITQEVLAACGRPOATAVYKYEYARRRGVPIVADGGIYONNGHIAK 375
DB 312 AADGGRIRMGSGSICITQEVLAACGRPOATAVYKYEYARRRGVPIVADGGIYONNGHIAK 371
OY 376 ALATGASTVMGSLAATTEARPEGEFFSDGIRLKYRKGSLDAMDKHLSSONRYFSEAD 435
DB 372 ALATGASTVMGSLAATTEARPEGEFFSDGIRLKYRKGSLDAMDKHLSSONRYFSEAD 431
OY 436 KIRVAQVGSAGVQDQGSIHKFPYLIAGIYHSCODIGAKSLTQVRAMVSGELKFEKRTS 495
DB 432 AVQVAGVGSAGVQDQGSIAKLIAYVSKGLQSAODIGELSPAIRREKMYAGQVLFPSRSP 491
OY 496 SAQVEGCVHSLHSTYERKLF 514
DB 492 TAQSEGCVHSLHSTYERKLF 510

```

## RESULT 12

A38668  
 IMP dehydrogenase (EC 1.1.1.205) - Leishmania donovani  
 N:Alternate names: IMPDH; inosine 5'-monophosphate dehydrog  
 C:Species: Leishmania donovani  
 C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_C  
 C:Accession: A38668  
 R:Wilson, K.; Collart, F.R.; Huberman, E.; Stringer, J.R.;  
 J. Biol. Chem. 266, 1665-1671, 1991  
 A:Title: Amplification and molecular cloning of the IMP de  
 A:Reference number: A38668; MUID:91107664; PMID:1671039  
 A:Accession: A38668  
 A:Molecule type: DNA  
 A:Residues: 1-514 <MIL>  
 A:Cross-references: GB:M55667; NID:9159360  
 C:Function:  
 A:Description: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecule of

A:Pathway: purine nucleotide biosynthesis  
 C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homo  
 C:Keywords: GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosynthesis  
 F:28-94/Domain: IMP dehydrogenase amino-terminal homology <IDHN>  
 F:179-228/Domain: CBS homology <CBS>  
 F:229-472/Domain: IMP dehydrogenase catalytic homology <IDHC>  
 F:327/Active site: Cys #status predicted

Query Match 52.1%; Score 1363.5; DB 1; Length 514;  
 Best Local Similarity 53.5%; Pred. No. 1.3e-80;  
 Matches 267; Conservative 83; Mismatches 146; Indels 3; Gaps 2;

```

OY 16 DGLTAQOLNCGDGLTYNDFLIPGYIDPTAOVDLTSLTKRKLTKLTLPVSSPMDTTE 75
DB 15 DCGTAEELFR-ODGLTYNDFILPGFIDPDGADVNDISQGFTRIRLHPTVSSPMDTTE 73
OY 76 AGMAIAMLTGIGIFHNCTPEFOANEVRYKRYEKGFTDPVYLSPKDRVDFEAKA 135
DB 74 NEAKTAMALMGVGLVHNCTVERQEVNYSKAYRNGFTSKPKSVPTPTISNIRIKE 133
OY 136 RHGFCGIPITDTRGMSRLVGISSNDIDFLKEEHDCLFEIMTKREDLVAPRSITLK 195
DB 134 EKGISGILVTENGDPHKGILGIVCTKIDIDYVANK--DTPVSAVMTRREKMYERAPITOLE 191
OY 196 EANEILORSKKGKLPVNEDEDLVALIARTDLKKNRDYPLASKDAKKOLLGCAAGTCHED 255
DB 192 EAMVDLNRSGRYGLPVNENDEVNLCSRDVARRADYPSHTLDSGRILCAAASTREPE 251
OY 256 DKRYRLDLAQAQVDVYVLDSSQGNSTFQINMTIKYKDKPNIQVIGNVYTAQAOKNLID 315
DB 252 DKRYRLDLAQAQVDVYVLDSSQGNSTFQINMTIKYKDKPNIQVIGNVYTAQAOKNLID 311
OY 316 AGVDALRVMGSGSICITQEVLAACGRPOATAVYKYEYARRRGVPIVADGGIYONNGHIAK 375
DB 312 AADGGRIRMGSGSICITQEVLAACGRPOATAVYKYEYARRRGVPIVADGGIYONNGHIAK 371
OY 376 ALATGASTVMGSLAATTEARPEGEFFSDGIRLKYRKGSLDAMDKHLSSONRYFSEAD 435
DB 372 ALATGASTVMGSLAATTEARPEGEFFSDGIRLKYRKGSLDAMDKHLSSONRYFSEAD 431
OY 436 KIRVAQVGSAGVQDQGSIHKFPYLIAGIYHSCODIGAKSLTQVRAMVSGELKFEKRTS 495
DB 432 AVQVAGVGSAGVQDQGSIAKLIAYVSKGLQSAODIGELSPAIRREKMYAGQVLFPSRSP 491
OY 496 SAQVEGCVHSLHSTYERKLF 514
DB 492 TAQSEGCVHSLHSTYERKLF 510

```

## RESULT 13

T32709  
 IMP dehydrogenase (EC 1.1.1.205) T322D1.3 - Caenorhabditis elegans  
 N:Alternate names: IMP; NAD+ oxidoreductase; IMPDH; Inosine-5'-monophosphate dehydrog  
 C:Species: Caenorhabditis elegans  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C:Accession: T32709  
 R:Geisels, C.; Bradshaw, H.; Hawkins, M.  
 submitted to the EMBL Data Library, December 1997  
 A:Description: The sequence of C. elegans cosmid T322D1.  
 A:Reference number: Z31211  
 A:Accession: T32709  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-499 <GEI>  
 A:Cross-references: EMBL:AF039052; PIDN:AAB94282.1; GSPDB:GN00022; CESP:T322D1.3  
 A:Experimental source: strain Bristol N2; clone T322D1  
 C:Genetics:  
 A:Gene: CESP:T322D1.3  
 A:Map position: 4  
 A:Insertions: 35/2; 85/3; 159/1; 183/3; 342/1; 412/3; 493/2  
 C:Function:  
 A:Description: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecule  
 A:Pathway: purine nucleotide biosynthesis

C:Superfamily:IMP dehydrogenase: CBS homology: IMP dehydrogenase amino-terminal homology  
C:Keywords: duplication, GMP biosynthesis, NAD, oxidoreductase, purine nucleotide biosynthesis  
E:352-98/Domain: IMP dehydrogenase amino-terminal homology <IDHn>  
E:120-170/Domain: CBS homology <CBS1>  
E:186-238/Domain: CBS homology <CBS2>  
E:239-461/Domain: IMP dehydrogenase catalytic homology <IDHC>  
E:337/Active site: Cys #status predicted

Query Match	50.6%;	Score 1324.5;	DB 1;	Length 499;
Best Local Similarity	52.1%;	Pred. No. 4.2e-78;		
Matches 264;	Conservative 95;	Mismatches 115;	Indels 33;	Gaps 6;

QY 16 DGLTADOLFNCNGDGLTYNDELILPGYIDFTADQVVLTSALTKILTKPLVSSPMDDYTE 75  
 18 DEELYHEHMAHRAKAGLTLYNDENILPGFINFGVHDSLETRITKDLKAPLVSSPMDDYTE 77  
 QY 76 AGMATAMALTTGGIGFIHNC-TRFEQAVNEVRKKVYKEDGFTDPVYLSPKDRVRDVEFAK 134  
 78 SGMAIVMALYGGITGIHNEFRKPEDQAAEVLKRFKFGQYVQAPCLSRDSTAFPMIDIK 137  
 Db

QY 135 ARHGFCEIPTDGRGMSRLVGISSRDIDPLKEE---EHDCLFEIINTKREDLVNAPRS 191

Db 138 KRYGTGAVPTDEGRVSGSKLIIGVTSRDFEITMDVAGCGCTPSIDINVSQÖLHLC--- 194

[illegible]

OY 248 AALGTHEDDKYRLDLLAAGVGVVVVLDDSSQGSNFIQIMNMIKYIKDKPEYNLQVIGANNVTA 307  
||| : | : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : : |||||  
DB 254 AAIVNTGESGYTVDRVEGEVDPLITIDSSNGSSTOISMTRYIEKHPHVOYTAGNVNTR 313

[illegible][illegible]

QY 428 NRYFSEADKIRVAGVSGAVODKSIHKFPYLLIAGIOHSDCODIGAKSLTVRAMMYSGE 487  
|||:||||||| :||:| | | | | | | | | | :| :|  
416 --FESDQIKVAGVSGATMKRCKHETPYLLIRGVHOMDQIVRSIRDFEEKVNGT 472

0Y 488 LKEKRTSSAOVEGGVHSLHSYEKRLP 514  
| | | : : : | | | | | | | | |  
473 VKFFRRSTNDOLFEGGVHSLHSSEKPLV 499

## RESULT 14

C:Accession: AF6306  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:MP dehydrogenase (EC 1.1.1.205) - Arabidopsis thaliana

R.Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kall, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L. *Mathematics* **2020**, *8*, 2000.

A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.S.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Luoro, Z.A.; Luoro, J.S.; Malti, R.; Marzalli, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

```
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <STO>
```

C:\Genetics: C:\Genetics\17  
A:\Map position: 1

C: Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homology; C: Keywords: GMP biosynthesis; NAD; oxidoreductase

Query March	44.1%;	Score 1155;	DB 2;	Length 502;
Best Local Similarity	47.4%;	Pred. No. 3.9e-67;		
Matches 239;	Conservative 93;	Mismatches 162;	Indels 10;	Gaps 6

[illegible]

OY 75 EAGAIAMALTCGIGFIHNNCTEFOANEVRKAKKYEOGFITDPEVLSKDRPRVDVEFK 134  
+ : : : : :  
Db 65 ESHVAANAALGIGGIIVNCDIDTOASVIIRHAKSLQVPIASAVFCKCEHOHSI - DDE 123  
+ : : : : :

Qy 135 ARHFGCIPITDTGRMGSLVGLIISSRDIIDFLKEEHHDFLEIIMTKREDL-VVAPRSIT 193  
 124 GPSSF--VEVSOTGLTLFKLLGVYKSEWSMMDDOKEVKIDYPMKSCNNKDYVPMWID 181

Oy 194 LKAEANIILÖRSKKGLPIVNEDEDLVAIARTDLKNRDP--LASKDAKKOLLGCAAI 250  
+ : : : : + : : : : + : : : :  
Db 182 IJDI FAIVLEFDKRG - FVYIEKEFGTGVNVTKDIVERVKYEPKIGSGTGACDKKMVGAI 240

[illegible][illegible]

OY 371 GHAKALALGASTVMGSLAATFEAPGEFFESDGI RLKKYRGSGLDAMDKHLSSONRY 430  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 361 GHTVKAIVTCASTVMGSELTAGSPFADGAFFEVNCPKPRKYRPGMSFTMTR--GSTDDY A18

OY 431 FSEEDRIKAVOGSGA VODKSTHKEFVPLIAGIOHSCODIGASLTQVRAMYSGEIKF 490  
 : | : ||| | | | | | : : : | | :  
 418 ICEBAKTIKOCUCVADKDCU KFEIVEMWAKGECEDICESTIOCHETIPDNETPI 478

QY 491 EKRSSAOEGVHSLHSYERLRF 514  
+ | | : | | : | | | | |  
DB 479 EAPRCAIPIKCTIICIVYERKST 503

## RESULT 15

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

R:collart, F., Osipuk, J., Trent, J., Olsen, G.J., Huberman, E.  
Gene 174, 217-220. 1996  
A:Title: Cloning and characterization of the gene encoding IMP dehydrogenase from *Neisseria meningitidis*

A:Accession: JC4999  
A:Molecule type: DNA  
A:Residues: 1-503 <COL>

C:genetics:  
A:gene: impdh  
A:introns: 135/1; 334/3; 404/3; 490/3

C:Function: provides precursors for DNA and RNA biosynthesis; it catalyzes the conversion of 5-methylthioadenosine to 5-methylthioadenosine diphosphate  
A:Description: provides precursors for DNA and RNA biosynthesis; it catalyzes the conversion of 5-methylthioadenosine to 5-methylthioadenosine diphosphate  
C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal domain

E:20-86/Domain: 1MP dehydrogenase amino-terminal homology <IDHN>  
E:20-86/Domain: 1MP dehydrogenase catalytic homology <IMD>  
F:172-220/Domain: CBS homology <CBS2>  
F:221-465/Domain: IMP dehydrogenase catalytic homology <IMD>

Query Match	42.28;	Score 1106.5;	DB 1;	Length 503;
Best Local Similarity	47.28;	Pred. No. 5.3e-64;		

Query Match	42.28;	Score 1106.5;	DB 1;	Length 503;
Best Local Similarity	47.28;	Pred. No. 5.3e-64;		





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 20:31:00 : Search time 209 Seconds

(without alignment updates/sec  
11436.455 Million cell

Title: US-09-846-637c-3

Perfect score: 1654

Sequence: 1 gaattcgagcgctctcgcga.....gttagaagaccgaattc 1654

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 41/9280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

- 1: gb.ba:\*
- 2: gb.htg:\*
- 3: gb.in:\*
- 4: gb.om:\*
- 5: gb.ov:\*
- 6: gb.pat:\*
- 7: gb.ph:\*
- 8: gb.pl:\*
- 9: gb.pr:\*
- 10: gb.ro:\*
- 11: gb.sts:\*
- 12: gb.sy:\*
- 13: gb.un:\*
- 14: gb.vl:\*
- 15: em.ba:\*
- 16: em.fun:\*
- 17: em.hum:\*
- 18: em.in:\*
- 19: em.mu:\*
- 20: em.om:\*
- 21: em.or:\*
- 22: em.ov:\*
- 23: em.pat:\*
- 24: em.ph:\*
- 25: em.pl:\*
- 26: em.ro:\*
- 27: em.sts:\*
- 28: em.un:\*
- 29: em.vl:\*
- 30: em.htg.hum:\*
- 31: em.htg.inv:\*
- 32: em.htg.other:\*
- 33: em.htg.mus:\*
- 34: em.htg.pln:\*
- 35: em.htg.tod:\*
- 36: em.htg.mam:\*
- 37: em.htg.vrt:\*
- 38: em.sy:\*
- 39: em.htgo.hum:\*
- 40: em.htgo.mus:\*
- 41: em.htgo.other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	87.7	1654	9 HUMIMP	J04208 Human inosi
2	1438	86.9	1642	6 ARI41889	ARI41889 Sequence
3	1438	86.9	1642	6 106570	106570 Sequence 1
4	1178	71.2	1663	9 BC006124	BC006124 Homo sapi
5	1172	70.9	1648	9 BC012840	BC012840 Homo sapi
6	1170	70.7	1655	9 BC015567	BC015567 Homo sapi
7	615	37.2	1158	6 ARI418035	ARI418035 Sequence
8	613	37.1	1158	6 ARI418036	ARI418036 Sequence
9	612	37.0	1155	6 ARI418030	ARI418030 Sequence
10	612	37.0	1155	6 ARI418031	ARI418031 Sequence
11	612	37.0	1155	6 ARI418032	ARI418032 Sequence
12	612	37.0	1155	6 ARI418033	ARI418033 Sequence
13	612	37.0	1158	6 ARI418037	ARI418037 Sequence
14	396	23.9	396	6 AX093346	AX093346 Sequence
15	242	14.6	6193	6 AX095550	AX095550 Sequence
16	242	14.6	6193	9 HUMIMP	L33842 Homo sapien
17	209	12.6	10537	9 HUMIMP	L33210 Homo sapien
18	137	8.3	1851	9 BC009321	BC009321 Homo sapi
19	88	5.3	360	9 HUMIMP	L08114 Human inosi
20	62	3.7	1614	10 CRUMP	J04209 Chinese ham
21	62	3.7	1620	6 ARI41890	ARI41890 Sequence
22	62	3.7	1620	6 106572	106572 Sequence 3
23	56	3.4	1545	10 MUSIMP	M88333 Mus musculu
24	56	3.4	1680	10 BC010314	BC010314 Mus muscul
25	56	3.4	1707	10 MUSIMP	M33934 Mouse IMP d
26	56	3.4	182658	2 AC084743	AC084743 Mus muscu
27	56	3.4	225727	2 AC122514	AC122514 Mus muscu
28	53	3.2	332	9 HS86C2R	Z66268 H.sapiens C
29	51	3.1	51	6 AX190186	AX190186 Sequence
30	41	2.5	393	6 ARI41893	ARI41893 Sequence
31	41	2.5	393	6 ARI41894	ARI41894 Sequence
32	41	2.5	393	6 106574	106574 Sequence 5
33	39	2.4	114558	2 AC107280	AC107280 Rattus no
34	39	2.4	16031	2 AC129612	AC129612 Rattus no
35	38	2.3	246	10 MMU89404	MMU89404 Mus muscul
36	33	2.0	168808	2 AC103220	AC103220 Rattus no
37	30	1.8	44	6 ARI19642	ARI19642 Sequence
38	29	1.8	687	9 HUMYO19A04	AF075074 Homo sapi
39	28	1.7	6193	6 AX348472	AX348472 Sequence
40	28	1.7	114558	2 AC107280	AC107280 Rattus no
41	26	1.6	42	6 ARI19643	ARI19643 Sequence
42	26	1.6	50	6 AX190187	AX190187 Sequence
43	26	1.6	51	6 AX190185	AX190185 Sequence
44	25	1.5	35	6 AX418043	AX418043 Sequence
45	25	1.5	51	6 AX190184	AX190184 Sequence
46	24	1.5	33	6 AX418042	AX418042 Sequence
47	23	1.4	30	6 AX418045	AX418045 Sequence
48	23	1.4	344	6 ARI41891	ARI41891 Sequence
49	23	1.4	344	6 ARI41892	ARI41892 Sequence
50	23	1.4	344	6 106573	106573 Sequence 4

ALIGNMENTS

RESULT 1  
HUMIMP 1654 bp mRNA linear PRI 11-JUN-1993  
LOCUS Human inosine-5'-monophosphate dehydrogenase (IMP) mRNA, complete cds.  
J04208  
ACCESSION J04208.1 GI:186391  
VERSION inosine-5'-monophosphate dehydrogenase.  
KEYWORDS Human peripheral blood leukocyte, cDNA to mRNA, clone HIMP.  
SOURCE Homo sapiens  
ORGANISM

Pred. No. is the number of results predicted by change to have a

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1654)  
AUTHORS  
Collart, F.R. and Huberman, E.  
TITLE  
Cloning and sequence analysis of the human and Chinese hamster  
inosine-5'-monophosphate dehydrogenase cDNAs  
J. Biol. Chem. 263 (30), 15769-15772 (1988)  
JOURNAL  
9008491  
MEDLINE  
2902093  
PUBMED

COMMENT  
Draft entry and computer-readable sequence for [1] kindly provided  
by F.R.Collart, 22-Aug-1988.

FEATURES  
Location/Qualifiers

source

CDS

1..1654  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
48..1592  
/note="inosine-5'-monophosphate dehydrogenase (EC  
1.1.1.205)"  
/codon\_start=1  
/protein\_id="AAA36112.1"  
/translation="MADYLIISGTSYVPDDGLTAQQLFNGDGLTYNDLILPGYIDF  
TADVDLSALTIKKILTKPLVSSPMQVTEAGMAIMALATIGTGFHNCTPEPOAN  
EVRKYKKEGGFTIDPVVLSPKDRDYFEKARHGFGLIPIITDGRKSLVGLTIS  
RDIDPLKEEEDCELEITMRREDLVAPRSITLKEANEILQRSKGLPLVNEDEL  
VAIARNDLKNRNDYPLASDKAKQLGGAIGHEDDKYLRQAGVQVYVDDSS  
QGNSTFQINMKYIKDKPNDQVIGNVNVAQAALDAGDALRYMGSSGICITQ  
EVLAGEPQATVYKRYSEYARFQVPLADGILQVHIALALAGASTVMGSLAA  
TTPEAGEFQSDGIRLKKYRGMGLDADKHLSSONRREFSADRIKVNQVGYAODK  
GSIKFPYLLAGIOLHSCODIGAKSLITVYRAMTSGELKEFKRTSSAOVBGVSHLS  
YEKRLF"

BASE COUNT 415 a 418 c 451 g 370 t  
ORIGIN 45 bp upstream of Bali site.

Query Match 87.7%; Score 1450; DB 9; Length 1654;  
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1650; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAATTCGGGGGCTCTCTCGAGACACCGCGGCTGTCCTGTTGGCCATGGCCGATACC 60  
DB 1 GAATTCGGGGGCTCTCTCGAGACACCGCGGCTGTCCTGTTGGCCATGGCCGATACC 60  
QY 61 TGATTAGTGGGGGACGCTCTACGTGCGACAGAGAGGACTCACAGACAGCAGCTTTCA 120  
DB 61 TGATTAGTGGGGGACGCTCTACGTGCGACAGAGAGGACTCACAGACAGCAGCTTTCA 120  
QY 121 ACTGGGGAGAGGCGCTCACTACAAATGACTTTCATCTCCCTGGGTCATCGACTTCA 180  
DB 121 ACTGGGGAGAGGCGCTCACTACAAATGACTTTCATCTCCCTGGGTCATCGACTTCA 180  
QY 181 CTGCAGACCAAGGTGACCTGACTTGTCTGTACCAAGAAATCACTTTAAGACCCAC 240  
DB 181 CTGCAGACCAAGGTGACCTGACTTGTCTGTACCAAGAAATCACTTTAAGACCCAC 240  
QY 241 TGGTTCTCTCTCCATGACACAGTACACAGAGGCTGGATGGCCATAGCAATGGCGCTTA 300  
DB 241 TGGTTCTCTCTCCATGACACAGTACACAGAGGCTGGATGGCCATAGCAATGGCGCTTA 300  
QY 301 CAGGGGGTFTTGGCTTCACTACCACTGTAACACTTAATTCAGGGCCATGAAGTTC 360  
DB 301 CAGGGGGTFTTGGCTTCACTACCACTGTAACACTTAATTCAGGGCCATGAAGTTC 360  
QY 361 GGAAGTGAAGAAATATGAACAGAGGATTCATCAGAGCCCTGTGGTCTCAGCCCCAAG 420  
DB 361 GGAAGTGAAGAAATATGAACAGAGGATTCATCAGAGCCCTGTGGTCTCAGCCCCAAG 420  
QY 421 ATCGCGTGGGAGATGTTTGAAGGCCAAGGCCCGGATGTTTGGCGTATCCCAATCA 480  
DB 421 ATCGCGTGGGAGATGTTTGAAGGCCAAGGCCCGGATGTTTGGCGTATCCCAATCA 480  
QY 481 CAGACACAGGCGGATGGGGAGCGGCTTGGTGGGCTCATCTCTCCAGGACATTGATT 540  
DB 481 CAGACACAGGCGGATGGGGAGCGGCTTGGTGGGCTCATCTCTCCAGGACATTGATT 540

QY 541 TTCTCAAGAGGAGAGACATGACTTTTCTTGGAAAGATAATGACAAAGAGGAACT 600  
DB 541 TTCTCAAGAGGAGAGACATGACTTTTCTTGGAAAGATAATGACAAAGAGGAACT 600  
QY 601 TGGTGTAGCCCCCGGACGATCACACTGAGAGGCAATTAATTTCTGCGAGCCAGCA 660  
DB 601 TGGTGTAGCCCCCGGACGATCACACTGAGAGGCAATTAATTTCTGCGAGCCAGCA 660  
QY 661 AAGAGGAAAGTTGCCCATTTGAATGAAGATGATGAGCTTGGCCATCATTTGCCGGA 720  
DB 661 AAGAGGAAAGTTGCCCATTTGAATGAAGATGATGAGCTTGGCCATCATTTGCCGGA 720  
QY 721 CAGACCTGAAGAAAGATCGGAGCTACCACTAGCCTCCAAAGATGCCAAGAAAGCTGC 780  
DB 721 CAGACCTGAAGAAAGATCGGAGCTACCACTAGCCTCCAAAGATGCCAAGAAAGCTGC 780  
QY 781 TGTGTGGGCGACCATTTGGCACTCATGAGGATGACAAGATATAGGCTGGACTTCCGCC 840  
DB 781 TGTGTGGGCGACCATTTGGCACTCATGAGGATGACAAGATATAGGCTGGACTTCCGCC 840  
QY 841 AGGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 AGGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 AATGATCAAGTACATCAAGCAAAATACCTTAATCTCCAAATTTGAGGCAATGTGG 960  
DB 901 AATGATCAAGTACATCAAGCAAAATACCTTAATCTCCAAATTTGAGGCAATGTGG 960  
QY 961 TCACCTGTCTCCAGGCGCAAGAACCTTCATTTGATGACAGGTGTGATGATGATGATGAT 1020  
DB 961 TCACCTGTCTCCAGGCGCAAGAACCTTCATTTGATGACAGGTGTGATGATGATGATGAT 1020  
QY 1021 TGGGAAGTGTCTCCATCTGATATTCACAGAGTGTGATGATGATGATGATGATGATGAT 1080  
DB 1021 TGGGAAGTGTCTCCATCTGATATTCACAGAGTGTGATGATGATGATGATGATGATGAT 1080  
QY 1081 CAGCAGTGTACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1081 CAGCAGTGTACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
QY 1141 GAGGAATCCAAATATGGGCTCATTTGGCAAGGCTTGGGCTTGGGCTTCCACAGTCA 1200  
DB 1141 GAGGAATCCAAATATGGGCTCATTTGGCAAGGCTTGGGCTTGGGCTTCCACAGTCA 1200  
QY 1201 TGTATGAGCTCTCTCTCGGCTGACACACTGAGAGGCGCTGTGATATCTTTCCGATG 1260  
DB 1201 TGTATGAGCTCTCTCTCGGCTGACACACTGAGAGGCGCTGTGATATCTTTCCGATG 1260  
QY 1261 GGAATCGGCTAAAGAAATATGCGGCTATGAGTCTCTCGATGCCATGACAAAGACCTCA 1320  
DB 1261 GGAATCGGCTAAAGAAATATGCGGCTATGAGTCTCTCGATGCCATGACAAAGACCTCA 1320  
QY 1321 GCAGCCAGAACGATATTTCACTGAAAGCTGACAAATAATCAATGAGGCCCGAGAGTGTCTG 1380  
DB 1321 GCAGCCAGAACGATATTTCACTGAAAGCTGACAAATAATCAATGAGGCCCGAGAGTGTCTG 1380  
QY 1381 GTCGTGTGAGGACAAAGGCTCAATCCAAATTTGCTTACCTGATGCTGCGATCC 1440  
DB 1381 GTCGTGTGAGGACAAAGGCTCAATCCAAATTTGCTTACCTGATGCTGCGATCC 1440  
QY 1441 AACACTCATGCCAGACATTTGTCACAGAGCTTACCCAAAGTCCGAGGACATGATGTACT 1500  
DB 1441 AACACTCATGCCAGACATTTGTCACAGAGCTTACCCAAAGTCCGAGGACATGATGTACT 1500  
QY 1501 CTGGGAGCTTAAAGTTTGAAGAGACAGCTCTCAGCCAGTGGAAAGTGGCGTCCATA 1560  
DB 1501 CTGGGAGCTTAAAGTTTGAAGAGACAGCTCTCAGCCAGTGGAAAGTGGCGTCCATA 1560  
QY 1561 GCTTCATTCGATGAGAGGCGGCTTTCTGAAAGGAGTCCAGACACCTCTCGGTTT 1620  
DB 1561 GCTTCATTCGATGAGAGGCGGCTTTCTGAAAGGAGTCCAGACACCTCTCGGTTT 1620

Qy 1621 TTTTTCATTAAGTTTAGAAGACCCGAATTC 1654  
|||||  
Db 1621 TTTTTCATTAAGTTTAGAAGACCCGAATTC 1654

RESULT 2  
AR141889  
LOCUS AR141889 1642 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 1 from patent US 6147194.  
ACCESSION AR141889  
VERSION AR141889.1 GI:15101405  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1642)  
AUTHORS Collart, F. R. and Huberman, E.  
TITLE Eukaryotic IMPDH polynucleotide and antibody compositions  
JOURNAL Patent: US 6147194-A 1 14-NOV-2000;  
FEATURES  
source 1. 1642  
Location/Qualifiers

BASE COUNT 411 a 416 c 449 g 366 t  
ORIGIN  
Query Match Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 GGGCGCTCCGCGAGACAGCGCGGCTGCTGTTGGCCATGCGCG  
1 GGGCGCTCCGCGAGACAGCGCGGCTGCTGTTGGCCATGCGCG  
Db 1 GGGCGCTCCGCGAGACAGCGCGGCTGCTGTTGGCCATGCGCG  
Qy 67 GTGGGGGCGACGTCTACGTGCGCAGAGAGGACTACAGACAGCAGC  
61 GTGGGGGCGACGTCTACGTGCGCAGAGAGGACTACAGACAGCAGC  
Db 127 GAGAGCGGCTCCACCTACAGTACTTCTCATCTCCCTGGGTACATCG  
121 GAGAGCGGCTCCACCTACAGTACTTCTCATCTCCCTGGGTACATCG  
Qy 187 ACCAGGTGACCTGACTTCTGCTGTGACCAAGAAATCACTCTTAAGA  
181 ACCAGGTGACCTGACTTCTGCTGTGACCAAGAAATCACTCTTAAGA  
Db 247 CCTCTCCCATGGACACAGTACAGAGGCTGGATGGCCATAGCAATGG  
241 CCTCTCCCATGGACACAGTACAGAGGCTGGATGGCCATAGCAATGG  
Qy 307 GTATTGGCTTCAATCCACCACTGATACCTGAATTCAGGCCAATG  
301 GTATTGGCTTCAATCCACCACTGATACCTGAATTCAGGCCAATG  
Db 367 TGAAGAAATATGACAGGATTCATCAGACCGCTGGTCCACAGCC  
361 TGAAGAAATATGACAGGATTCATCAGACCGCTGGTCCACAGCC  
Qy 427 TCGCGGATGTTTTAGAGCCAGGCCGCGCATGTTTCTGCGGATTC  
421 TCGCGGATGTTTTAGAGCCAGGCCGCGCATGTTTCTGCGGATTC  
Db 487 CAGGCGGATGTTGGAGCGCGCTGTGGGATCATCTCCCTCCAGGAGCA  
481 CAGGCGGATGTTGGAGCGCGCTGTGGGATCATCTCCCTCCAGGAGCA  
Qy 547 AAGAGAGAGACATGACTGTTCTTGGAAGAGATATGACAAAGAGG  
541 AAGAGAGAGACATGACTGTTCTTGGAAGAGATATGACAAAGAGG  
Db 607 TAGCCCCCGCGAGCATCACACTGAAGAGAGCAATGATTTCTGAGCG  
601 TAGCCCCCGCGAGCATCACACTGAAGAGAGCAATGATTTCTGAGCG

Qy 667 GAAAGTTGCCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCCGAGACACC  
661 GAAAGTTGCCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCCGAGACACC  
Db 727 TGAAGAGATGCGGGAGCTACCCACTAGCCTCCAAAGATGCCAAGAAACAGTGTGTG  
721 TGAAGAGATGCGGGAGCTACCCACTAGCCTCCAAAGATGCCAAGAAACAGTGTGTG  
Qy 787 GGGCAGCCATTTGGCACTCTGAGAGATGACAGAGTATAGCTGTGACTTGTGCGCCAGGCTG  
781 GGGCAGCCATTTGGCACTCTGAGAGATGACAGAGTATAGCTGTGACTTGTGCGCCAGGCTG  
Qy 847 GTGTGATGATGATGTTTTGGAAGTCTTCCAGAGAAATTCATCTTTCAGATCAATATGA  
841 GTGTGATGATGATGTTTTGGAAGTCTTCCAGAGAAATTCATCTTTCAGATCAATATGA  
Db 907 TCAAGTACATCAAAAGCAAAATACCTTAATCTCAATCTCAATGAGGCAATGTGTCTACG  
901 TCAAGTACATCAAAAGCAAAATACCTTAATCTCAATCTCAATGAGGCAATGTGTCTACG  
Qy 967 CTGCCAGGCCCAAGAACTCATTTGATGACAGTGTGATGCCCTGCGGGTGGCCATGGGAA  
961 CTGCCAGGCCCAAGAACTCATTTGATGACAGTGTGATGCCCTGCGGGTGGCCATGGGAA  
Db 1027 GTGGCTCCATCTGCAATATCCAGAGAGTGTGCTGGCGTGGGCGCCCAAGCAGCAG  
1021 GTGGCTCCATCTGCAATATCCAGAGAGTGTGCTGGCGTGGGCGCCCAAGCAGCAG  
Qy 1087 TGTACAGGTGTATGATGATGACAGCGCGCTTGTGCTGCTGCTGATGTGAGGAA  
1081 TGTACAGGTGTATGATGATGACAGCGCGCTTGTGCTGCTGCTGATGTGAGGAA  
Db 1147 TCCAAATGTGGGTATATTTGGAAGACCTTGGCCCTTCACAGTATGATGG  
1141 TCCAAATGTGGGTATATTTGGAAGACCTTGGCCCTTCACAGTATGATGG  
Qy 1207 GCTCTCTCTGGCTGACACACTGAGAGCCCTGGTGAATACCTTCTTTCGATGGATCC  
1201 GCTCTCTCTGGCTGACACACTGAGAGCCCTGGTGAATACCTTCTTTCGATGGATCC  
Qy 1267 GGTAAAGAAATATGCGGATGATGGTCTCTGATGCCATGACAGACCTCAGACGC  
1261 GGTAAAGAAATATGCGGATGATGGTCTCTGATGCCATGACAGACCTCAGACGC  
Db 1327 AGACAGATATTTAGTGAAGCTGACAAATCAAAATGAGGCCAGAGAGTGTGGTCTG  
1321 AGACAGATATTTAGTGAAGCTGACAAATCAAAATGAGGCCAGAGAGTGTGGTCTG  
Qy 1387 TGCAGGACAAAGGTCATATCCAAATTTGTCCCTTACCTGATGCTGATCCAACT  
1381 TGCAGGACAAAGGTCATATCCAAATTTGTCCCTTACCTGATGCTGATCCAACT  
Db 1447 CATGCCAGGACATTTGGTGCAGAGAGCTTGACCAAGTCCGAGCCATGATGCTCGGGG  
1441 CATGCCAGGACATTTGGTGCAGAGAGCTTGACCAAGTCCGAGCCATGATGCTCGGGG  
Qy 1507 AGCTTAAGTTTGAAGAGACAGCTCTCAGCCAGGTGGAAGGTGGCTGATAGCTCC  
1501 AGCTTAAGTTTGAAGAGACAGCTCTCAGCCAGGTGGAAGGTGGCTGATAGCTCC  
Db 1567 ATTCGATGAGAGAGGCGCTTCTGAAAGAGGATCAGACAGCTCTCGGTTTTTTTTT  
1561 ATTCGATGAGAGAGGCGCTTCTGAAAGAGGATCAGACAGCTCTCGGTTTTTTTTT  
Qy 1627 CAATAAAGTTTGAAGAGACC 1648  
1621 CAATAAAGTTTGAAGAGACC 1642

RESULT 3  
LOCUS 106570 1642 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 9001545.

ACCESSION 106570  
VERSION 106570.1 GI:589653  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1642)  
AUTHORS Collart,F.R. and Huberman,E.  
TITLE METHOD AND MATERIALS RELATING TO IMPDH AND GMP PRODUCTION  
JOURNAL Patent: WO 9001545-A 1 22-FEB-1990;  
FEATURES  
source 1..1642  
location/Qualifiers  
BASE COUNT 411 a 416 c 449 g 366 t  
ORIGIN  
Query Match 86.9%; Score 1438; DB 6; Length 1642;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 GGGCGGTCCGAGACAGCGGGGGTGCCTGTGTGGCCATGGCCGACTACTGATTA 66  
DB 1 GGGCGGTCTCGAGACAGCGGGGGTGTCTGTGTGGCCATGGCCGACTACTGATTA 60  
QY 67 GTGGGGGCAAGTCTCTACGTGCGAGACGAGGACTACAGACAGAGCTCTTCACTGCG 126  
DB 61 GTGGGGGCAAGTCTCTACGTGCGAGAGGAGGACTACAGACAGAGCTCTTCACTGCG 120  
QY 127 GAGACGGCTCAGCTTCAATGACTTTCTCTCTCCCTGGGTACATGACTTCTGACG 186  
DB 121 GAGACGGCTCAGCTTCAATGACTTTCTCTCTCCCTGGGTACATGACTTCTGACG 180  
QY 187 ACCAGGTGAGCTGACTTCTGCTGTACCAAGAAATCACTCTTAAGACCCCTGTTT 246  
DB 181 ACCAGGTGAGCTGACTTCTGCTGTACCAAGAAATCACTCTTAAGACCCCTGTTT 240  
QY 247 CCTCTCCATGAGACAGTACAGAGGCTGGATGAGCATAGCAATGGCGTTTACAGCG 306  
DB 241 CCTCTCCATGAGACAGTACAGAGGCTGGATGAGCATAGCAATGGCGTTTACAGCG 300  
QY 307 GTATTGGCTTCACTCCACCAACTGTATCAACTGATTTCCAGGCCATGAAGTTCCGAAG 366  
DB 301 GTATTGGCTTCACTCCACCAACTGTATCAACTGATTTCCAGGCCATGAAGTTCCGAAG 360  
QY 367 TGAAGAAATATGAACAGGAGATTTCATACAGACCTGTGTCACAGCCCAAGAGATCGCG 426  
DB 361 TGAAGAAATATGAACAGGAGATTTCATACAGACCTGTGTCACAGCCCAAGAGATCGCG 420  
QY 427 TCGGGATGTTTTTGAAGCCAGGCGCGGATGTTTCTGCGGTATCCCAATCACAGACA 486  
DB 421 TCGGGATGTTTTTGAAGCCAGGCGCGGATGTTTCTGCGGTATCCCAATCACAGACA 480  
QY 487 CAGGGCGGATGGGGAGCGGCTGTGGGGATCATCTCCACAGGACATTTGATTTCTCA 546  
DB 481 CAGGGCGGATGGGGAGCGGCTGTGGGGATCATCTCCACAGGACATTTGATTTCTCA 540  
QY 547 AAGAGAGAGAACATGACTTTCTTCTTGAAGAGATATGACAAAGAGAGGAAGACTTGTCG 606  
DB 541 AAGAGAGAGAACATGACTTTCTTCTTGAAGAGATATGACAAAGAGAGGAAGACTTGTCG 600  
QY 607 TAGCCCCCGCAGCATACACTGAGAGGCAAAATTTCTGACGCGCAGCAAGAGG 666  
DB 601 TAGCCCCCGCAGCATACACTGAGAGGCAAAATTTCTGACGCGCAGCAAGAGG 660  
QY 667 GAAAGTTGCCCATTTGAATGAAGATGATGAGTTTGGCCATCATTTGCCCGACAGACC 726  
DB 661 GAAAGTTGCCCATTTGAATGAAGATGATGAGTTTGGCCATCATTTGCCCGACAGACC 720  
QY 727 TGAAGAAAGATGGGAGTACCCACTAGCCCTCCAAAGATGCAAGAAAGAGTCTGTG 786  
DB 721 TGAAGAAAGATGGGAGTACCCACTAGCCCTCCAAAGATGCAAGAAAGAGTCTGTG 780  
QY 787 GGGCAGCATTTGCACTCATGAGGATGACAAATATAGGCTGCACTTCTGCGCCAGGCTG 846

DB 781 GGGCAGCATTTGCACTCATGAGGATGACAAATATAGGCTGCACTTCTGCGCCAGGCTG 840  
QY 847 GTGTGATGTAGTGGTTTGGACTCTTCCAGGAAATTCATTTCCAGATCAATATGA 906  
DB 841 GTGTGATGTAGTGGTTTGGACTCTTCCAGGAAATTCATTTCCAGATCAATATGA 900  
QY 907 TCAAGTACATCAAGACAATACCCATATCTCCAAGTCAATTGAGGCAATGTGTCTACG 966  
DB 901 TCAAGTACATCAAGACAATACCCATATCTCCAAGTCAATTGAGGCAATGTGTCTACG 960  
QY 967 CTGCCCCAGCCAAAGACCTCATTTGATGACAGGTGTGAGTCCCTGCGGGTGGCATGCGAA 1026  
DB 961 CTGCCCCAGCCAAAGACCTCATTTGATGACAGGTGTGAGTCCCTGCGGGTGGCATGCGAA 1020  
QY 1027 GTGGCTCCATTCGATTTATCCAGGAAGTCTGCGCTGTGGGGGCCCCCAAGCAACAGCAG 1086  
DB 1021 GTGGCTCCATTCGATTTATCCAGGAAGTCTGCGCTGTGGGGGCCCCCAAGCAACAGCAG 1080  
QY 1087 TGTACAAAGTGTATGATGTGACAGGCGGCTTGTGTTCGGTCAATTGCTGATGAGGAA 1146  
DB 1081 TGTACAAAGTGTATGATGTGACAGGCGGCTTGTGTTCGGTCAATTGCTGATGAGGAA 1140  
QY 1147 TCCAAATGTGGGTATATTTGGGAAAGCCCTTGAGCCCTTGAGGCTTCACAGTCAATGATG 1206  
DB 1141 TCCAAATGTGGGTATATTTGGGAAAGCCCTTGAGCCCTTGAGGCTTCACAGTCAATGATG 1200  
QY 1207 GCTCTCTCTGCTGCTGCGACACCTAGAGGCCCCGTGTGTAATACTTTCTTTCCGATGGGATGC 1266  
DB 1201 GCTCTCTCTGCTGCTGCGACACCTAGAGGCCCCGTGTGTAATACTTTCTTTCCGATGGGATGC 1260  
QY 1267 GGTAAAGAAATATGCGGATATGGGTTCTCTGATGCCATGGACAAGCACTCAGACGCC 1326  
DB 1261 GGTAAAGAAATATGCGGATATGGGTTCTCTGATGCCATGGACAAGCACTCAGACGCC 1320  
QY 1327 AGAACAGATATTTCACTGTAAGTGTACAAAATCAAAAGTGCCCAAGGAGTGTCTGTGCTG 1386  
DB 1321 AGAACAGATATTTCACTGTAAGTGTACAAAATCAAAAGTGCCCAAGGAGTGTCTGTGCTG 1380  
QY 1387 TGCAGGCAAAAGGGTCAATCCCAAAATTTGTCCCTTACCTGATTTGTGTGATCCAAACT 1446  
DB 1381 TGCAGGCAAAAGGGTCAATCCCAAAATTTGTCCCTTACCTGATTTGTGTGATCCAAACT 1440  
QY 1447 CATGCCAGACATTTGGTGCAGAGGCTTGACCCAAATCCGAGCCATGATGTACTCTGGGG 1506  
DB 1441 CATGCCAGACATTTGGTGCAGAGGCTTGACCCAAATCCGAGCCATGATGTACTCTGGGG 1500  
QY 1507 AGCTTAAGTTTGAAGAGACAGTCTCTCAGCCAGGTGGAAGTGGCTTCATAGCTTCC 1566  
DB 1501 AGCTTAAGTTTGAAGAGACAGTCTCTCAGCCAGGTGGAAGTGGCTTCATAGCTTCC 1560  
QY 1567 ATTTCGTATGAGAAGCGGCTTTCTGTGAAGAGGATCCAGACACCTCTCGTGTGTTTTTTT 1626  
DB 1561 ATTTCGTATGAGAAGCGGCTTTCTGTGAAGAGGATCCAGACACCTCTCGTGTGTTTTTTT 1620  
QY 1627 CAATAAAGTTTGAAGAGACC 1648  
DB 1621 CAATAAAGTTTGAAGAGACC 1642

RESULT 4  
LOCUS BC006124 1663 bp mRNA linear PRI 12-JUL-2001  
DEFINITION Homo sapiens, IMP (inosine monophosphate) dehydrogenase 2, clone  
ACCESSION BC006124  
VERSION BC006124.1 GI:13543972  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1663)  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (02-Apr-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: cgarbs-remail.nih.gov  
Email: cgarbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada  
info@gscc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yarc Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Jarland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kusche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parveen Saeedi, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRML Plate: 18 Row: f Column: 15  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.  
location/Qualifiers  
1. 1663  
/organism="Homo sapiens"  
/db\_xref="locusID:3615"  
/db\_xref="taxon:9606"  
/clone="MGC:13021 IMAGE:3545693"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_lib="NIH-MGC\_15"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
41..1585  
/codon\_start=1  
/product="IMP (inosine monophosphate)"  
/protein\_id="AAH06124.1"  
/db\_xref="GI:13543973"  
/translacion="MADYLSGSGTSTVDPDGLTAOLFNCGDGLTYNDFLLIPGYIDE  
TADVDLTSLTKITTLKPLVSPMDYTEAGMAIALTGGIGFIHNNCPEROAN  
EVRKXKXEOGFIIDPVVLSPKDRVDFPEAKRGFGICIPITDGRMKSRLVFIOS  
RDIDLKEEHDCLEPIIMTKREDLVVAPAGITLKEAHEIIIDRSKKKPIRYNDEL  
VAIIRTDLKKRNDPLASKDKKOLLGCAIETHEDKYRIDLLAOKQDVVDLSS  
QONSIFQINMIKYIDKYPNLQVIGNVVTAQANILDAVDLALRGVSGSICITQ  
EVLACGRQATVAVYSEVAREFVPIADGIDQVGIADKALAGASVVMGSLIAA  
TTEARGEVEFFSDGIRLKRYRGMGSLDAMDKHLSSONRFSEADKIKVAGVGAQODK  
GSIHFVPIYLIAGIOHSCODIGAKSLITQVRAMVSGELKFEKRTSSAQVEGVSHLS  
YEKRLF"

BASE COUNT 436 a 412 c 448 g 367 t  
ORIGIN  
Query Match 71.2% Score 1178; DB 9; Length 1663;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1628; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 189 CAGGTGACCTGACTTCTGCTGACCAAGAAATACCTTAAGACCCACCTGTTCC 248  
|||||  
DB 182 CAGGTGACCTGACTTCTGCTGACCAAGAAATACCTTAAGACCCACCTGTTCC 241  
QY 249 TCTCCATGAGACACATCAGACAGAGCGTGGATGGCCATAGCAATGCGCTTACAGCGCT 308  
|||||  
DB 242 TCTCCATGAGACACATCAGACAGAGCGTGGATGGCCATAGCAATGCGCTTACAGCGCT 301  
QY 309 ATTGGCTTATCCACCAACCTGTACACCTGAATTCAGGCCAATGAAGTGGGAAGTG 368  
|||||  
DB 302 ATTGGCTTATCCACCAACCTGTACACCTGAATTCAGGCCAATGAAGTGGGAAGTG 361  
QY 369 AACAATATGAAACAGGATTCATCAGACACCGCTGCTCGCCGCCAAGATGCGCGT 428  
|||||  
DB 362 AACAATATGAAACAGGATTCATCAGACACCGCTGCTCGCCGCCAAGATGCGCGT 421  
QY 429 CGGGATGTTTGGAGCCCAAGGCCCGGCATGCTTCTGCGGTATCCCAATCAGACACA 488  
|||||  
DB 422 CGGGATGTTTGGAGCCCAAGGCCCGGCATGCTTCTGCGGTATCCCAATCAGACACA 481  
QY 489 GCGCGATGGGAGCGGCTTGGTGGGCATCATCTCCACAGGACATGATTTTCTCAA 548  
|||||  
DB 482 GCGCGATGGGAGCGGCTTGGTGGGCATCATCTCCACAGGACATGATTTTCTCAA 541  
QY 549 GAGGAGAACATGACTGTTTCTTGAAGATTAATGACAAAGAGGAGACTTGTGTA 608  
|||||  
DB 542 GAGGAGAACATGACTGTTTCTTGAAGATTAATGACAAAGAGGAGACTTGTGTA 601  
QY 609 GCGCCCGGACGATCAGCTGGAAGGAGGCAATGAATTCGAGGCGAGCAAGAGGA 668  
|||||  
DB 602 GCGCCCGGACGATCAGCTGGAAGGAGGCAATGAATTCGAGGCGAGCAAGAGGA 661  
QY 669 AAGTTCGCCATGTAATGAAGATGATGAGCTTGTGGCCATATTCGCCGACAGACTG 728  
|||||  
DB 662 AAGTTCGCCATGTAATGAAGATGATGAGCTTGTGGCCATATTCGCCGACAGACTG 721  
QY 729 AAGAAGATCGGACTACCACTAGCCCTCCCAAGATGACCAAGAAACAGCTGTGTGG 788  
|||||  
DB 722 AAGAAGATCGGACTACCACTAGCCCTCCCAAGATGACCAAGAAACAGCTGTGTGG 781  
QY 789 GCAGCATTGGGACATGAGGATGACAAATATAGCTGACTTGTGCGCCAGCGTGT 848  
|||||  
DB 782 GCAGCATTGGGACATGAGGATGACAAATATAGCTGACTTGTGCGCCAGCGTGT 841  
QY 849 GTGGATGTGTGGTGTGGACTCTTCCAGGGAATTCATCTCCAGATCAATATGATC 908  
|||||  
DB 842 GTGGATGTGTGGTGTGGACTCTTCCAGGGAATTCATCTCCAGATCAATATGATC 901  
QY 909 AAGTACATCAAGACAATACCTTAATCTCAAGTCAATGGAGCAATGTGCTACTGCT 968  
|||||  
DB 902 AAGTACATCAAGACAATACCTTAATCTCAAGTCAATGGAGCAATGTGCTACTGCT 961  
QY 969 GCGCCAGGCAAGACTCATGTATGACAGTGTGAGTGGCTCGGGTGGGCAATGGGAAT 1028  
|||||  
DB 962 GCGCCAGGCAAGACTCATGTATGACAGTGTGAGTGGCTCGGGTGGGCAATGGGAAT 1021  
QY 1029 GCGTCCATGCTGATTCAGGAAGAGTGTGAGCTGTGGGCGGCCCAACAGCAGTG 1088  
|||||  
DB 1022 GCGTCCATGCTGATTCAGGAAGAGTGTGAGCTGTGGGCGGCCCAACAGCAGTG 1081  
QY 1089 TACAAGGTATGAGTATCAGCGCGCTTGGTGGCTTCCGCTATTCGATGAGAGAAATC 1148  
|||||  
DB 1082 TACAAGGTATGAGTATCAGCGCGCTTGGTGGCTTCCGCTATTCGATGAGAGAAATC 1141  
QY 1149 CAAATGTGCGTCAATATTCGAAAGGCTTGGCCCTTGGGCGCTCAGACTCATGATGGCG 1208  
|||||  
DB 1142 CAAATGTGCGTCAATATTCGAAAGGCTTGGCCCTTGGGCGCTCAGACTCATGATGGCG 1201  
QY 1209 TCTCTCTGCGGTGCGACCACTGAGGCGCTGTGTGAATACTTCTTTCGATGGGATCGCG 1268  
|||||  
DB 1202 TCTCTCTGCGGTGCGACCACTGAGGCGCTGTGTGAATACTTCTTTCGATGGGATCGCG 1261

QY 1269 CTAAGAATATCGCGGTATGGGTTCTCTCGATGCCATGACAGACCTCAGCAGCCAG 1328  
 |||||||  
 Db 1262 CTAAGAATATCGCGGTATGGGTTCTCTCGATGCCATGACAGACCTCAGCAGCCAG 1321  
 |||||||  
 QY 1329 AACAGATATTTCGTGTAAGCTGCACAAATCAAGTGGCCAGGAGTGTCTGCTGTG 1388  
 |||||||  
 Db 1322 AACAGATATTTCGTGTAAGCTGCACAAATCAAGTGGCCAGGAGTGTCTGCTGTG 1381  
 |||||||  
 QY 1389 CAGGACAAAGGCGTCATCCACAAATTTGTCCCTTACCTGATGCTGCATCCAAACTCA 1448  
 |||||||  
 Db 1382 CAGGACAAAGGCGTCATCCACAAATTTGTCCCTTACCTGATGCTGCATCCAAACTCA 1441  
 |||||||  
 QY 1449 TCGCAGACATTTGCTGCCAAGAGCTTGACCCAGTCCAGCCATGATGATCTGTGGGAG 1508  
 |||||||  
 Db 1442 TCGCAGACATTTGCTGCCAAGAGCTTGACCCAGTCCAGCCATGATGATCTGTGGGAG 1501  
 |||||||  
 QY 1509 CTTAAGTTTGAAGAGACGTCCTCAGCCAGTGGAGTGGCTCCATACCTCCAT 1568  
 |||||||  
 Db 1502 CTTAAGTTTGAAGAGACGTCCTCAGCCAGTGGAGTGGCTCCATACCTCCAT 1561  
 |||||||  
 QY 1569 TCGTATGAGAGCGGCTTTCTGAAAAGGATCCAGACACACCTCTCGGTTTTTTTCA 1628  
 |||||||  
 Db 1562 TCGTATGAGAGCGGCTTTCTGAAAAGGATCCAGACACACCTCTCGGTTTTTTTCA 1621  
 |||||||  
 QY 1629 ATAAAAGTTTACAAGA 1645  
 |||||||  
 Db 1622 ATAAAAGTTTACAAGA 1638  
 |||||||

RESULT 5  
 LOCUS BC012840 1648 bp mRNA linear PRI 22-AUG-2001  
 DEFINITION Homo sapiens, IMP (inosine monophosphate) dehydrogenase 2, clone  
 GGC:9332 IMAGE:3447994, mRNA, complete cds.  
 ACCESSION BC012840  
 VERSION BC012840.1 GI:15277479  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1648)  
 Strausberg,R.  
 Direct Submission  
 Submitted (20-AUG-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.mci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalobobcm.tmc.edu](mailto:villalobobcm.tmc.edu)  
 Villalobon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Gibbs, R.A.

FEATURES  
 source  
 location/Qualifiers  
 1..1648  
 /organism="Homo sapiens"  
 /db\_xref="locusID:3615"  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRK Plate: 12 Row: a Column: 6  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein.

CDS  
 /db\_xref="taxon:9606"  
 /clone="MGC:9332 IMAGE:3447994"  
 /tissue\_type="Placenta, choriolealloma"  
 /clone\_id="NIH\_MGC\_10"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 34..1578  
 /codon\_start=1  
 /product="IMP (inosine monophosphate) dehydrogenase 2"  
 /protein\_id="AAH12840.1"  
 /db\_xref="GI:15277480"  
 /translation="MADYLISGTSYVPDGLTAAQLFNCGDGLTYNDLILPGYIDF  
 IADVDLSALTLLKTLTKPLVSPMDVTEAGMAIATLGGIGFIHNNCTPEROAN  
 EVKRYKKEGFIIDPVVLSPRDVRDVEAKARAGFCGIPITDGRMSRLVGLIS  
 RDIDLEKEEDHCCLEIEIMTRREDLVVAPAGITTEKEAEIILORSKKGKLPVNEDEL  
 VAIIRMDLKKNRDPLASKDKAKKOLLGCAITLEDDEKRYLLDLAOKGVVYVDS  
 OGNSIFQINMRYIKDKIPNIOVIGSNVTAQAKNLIDAGVDALRVMGSGSLITTO  
 EVLAGRPQATVTVKVEYARFVPAVDGILQNVGHIAALALAGASTVMGSLITTO  
 TTEARGEYFSDGIRLKRVRGMGLDAMDHLISQNRYSADAKIKVAGVGAVDK  
 GSIRHFVYLLAGIDHSCODIGAKSLTGVRAVMYSGELKFEKRTSSAOVEGVSLHS  
 YEKRLF"  
 BASE COUNT 428 a 409 c 445 g 366 t  
 ORIGIN

Query Match 70.9%; Score 1172; DB 9; Length 1648;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1622; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 15 CTCGGAGACACGCGGGGTCTCTGCTGTTGGCCATGCGCATCTACTGATTAAGTGGGCG 74  
 |||||||  
 Db 1 CTCGGAGACACGCGGGGTCTCTGCTGTTGGCCATGCGCATCTACTGATTAAGTGGGCG 60  
 |||||||  
 QY 75 ACCTCTAGCTCCAGACACAGGAGCTCACAGACAGACCTCTTCACTGCGGAGACGCG 134  
 |||||||  
 Db 61 ACCTCTAGCTCCAGACACAGGAGCTCACAGACAGACCTCTTCACTGCGGAGACGCG 120  
 |||||||  
 QY 135 CTCACCTACATGACTTTCTATCTCTCCCTGGGTACATGACTGACCTGACAGACGATG 194  
 |||||||  
 Db 121 CTCACCTACATGACTTTCTATCTCTCCCTGGGTACATGACTGACCTGACAGACGATG 180  
 |||||||  
 QY 195 GACCTGACTTTCGCTCTGACCAAGAAATCACTCTTAAGACCCCACTGGTTCTCTCC 254  
 |||||||  
 Db 181 GACCTGACTTTCGCTCTGACCAAGAAATCACTCTTAAGACCCCACTGGTTCTCTCC 240  
 |||||||  
 QY 255 ATGACACACGATCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGTATTGGC 314  
 |||||||  
 Db 241 ATGACACACGATCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGTATTGGC 300  
 |||||||  
 QY 315 TTCAATCCACCACTGTACACCTGTAATTCAGGCCAATGAAGTTGGAAGTGAAGAA 374  
 |||||||  
 Db 301 TTCAATCCACCACTGTACACCTGTAATTCAGGCCAATGAAGTTGGAAGTGAAGAA 360  
 |||||||  
 QY 375 TATGAACAGGATTCATACACACCTGTGTCCTACAGCCCAAGATTCGCTGGGGAT 434  
 |||||||  
 Db 361 TATGAACAGGATTCATACACACCTGTGTCCTACAGCCCAAGATTCGCTGGGGAT 420  
 |||||||  
 QY 435 GTTTTGAAGCCCAAGGCCCGGATGTTCTGCGGATCCCAATCACAGACACAGCGCGG 494  
 |||||||  
 Db 421 GTTTTGAAGCCCAAGGCCCGGATGTTCTGCGGATCCCAATCACAGACACAGCGCGG 480  
 |||||||  
 QY 495 ATGGGAGCGCGTGTGGTGGCATCTCTCCACAGGACATTTGATTTTCTCAAGAGAG 554  
 |||||||  
 Db 481 ATGGGAGCGCGTGTGGTGGCATCTCTCCACAGGACATTTGATTTTCTCAAGAGAG 540  
 |||||||  
 QY 555 GAACATGACTGTTTCTTGAAGAGATTAATGACAAAGAGGAAGACTTGGTGTAGCCCC 614  
 |||||||  
 Db 541 GAACATGACTGTTTCTTGAAGAGATTAATGACAAAGAGGAAGACTTGGTGTAGCCCC 600  
 |||||||  
 QY 615 CCGAGATCAACATGAGAGGCAATGAATTCGACGCGGAGAGAAAGAGGAAGTTG 674  
 |||||||  
 Db 601 CCGAGATCAACATGAGAGGCAATGAATTCGACGCGGAGAGAAAGAGGAAGTTG 660  
 |||||||  
 QY 675 CCAATTTGAATGAAGATGAGCTTGTGGCATCATTTGCCCGGACAGACTGAAGAAG 734  
 |||||||



```

Db      661  CCATTTGTAATGAAGATGATGAGCTTGTCGCATCATTCGCCGAGCA  BACCTGAAGAAG  720
Qy      735  AATCGGAGCTACCCACTAGCTTCCAAAGATGCCAGAAACAGCTGCTG  GTGGGGGAGCC  794
Db      721  AATCGGAGCTACCCACTAGCTTCCAAAGATGCCAGAAACAGCTGCTG  GTGGGGGAGCC  780
Qy      795  ATTGGCACTCATGAGATGACAAATATAGAGCTGCTGCTGCCAGC  CTGTGTGTGAT  854
Db      781  ATTGGCACTCATGAGATGACAAATATAGAGCTGCTGCTGCCAGC  CTGTGTGTGAT  840
Qy      855  GTAGTGGTTTGGAGCTCTTCCAGGGAATTCATCTCCAGATCAAT  ATGATCAAGTAC  914
Db      841  GTAGTGGTTTGGAGCTCTTCCAGGGAATTCATCTCCAGATCAAT  ATGATCAAGTAC  900
Qy      915  ATCAAGACAAATATACCTATCTCCAGATCATTCGAGCAATGTGTC  CTGTGTGTGAT  974
Db      901  ATCAAGACAAATATACCTATCTCCAGATCATTCGAGCAATGTGTC  CTGTGTGTGAT  960
Qy      975  GCCAAGAACCTCATGATGACAGTGTGATGCCCTGGGTGGCATG  GGAAGTGGCTCC  1034
Db      961  GCCAAGAACCTCATGATGACAGTGTGATGCCCTGGGTGGCATG  GGAAGTGGCTCC  1020
Qy      1035  ATCTGATATATCAGGAATGCTGGCTGTGGGGCCCCAAGCACA  ACAGTGTACAG  1094
Db      1021  ATCTGATATACGAGGAATGCTGGCTGTGGGGCCCCAAGCACA  ACAGTGTACAG  1080
Qy      1095  GTGTATGATATGACAGGCGCTTGTGTTCCGCTCATTCGTATGA  GGAATCCAAAT  1154
Db      1081  GTGTATGATATGACAGGCGCTTGTGTTCCGCTCATTCGTATGA  GGAATCCAAAT  1140
Qy      1155  GTGGGTCATATTCGGAACCTTGCCCTTGGGGCTCCACAGTCAAT  ATGGGCTCTTC  1214
Db      1141  GTGGGTCATATTCGGAACCTTGCCCTTGGGGCTCCACAGTCAAT  ATGGGCTCTTC  1200
Qy      1215  CTGGGCTGCCAGCACTGAGGCCCTTGTGTAATCTTCTTCGATG  GATTCGGCTTAAG  1274
Db      1201  CTGGGCTGCCAGCACTGAGGCCCTTGTGTAATCTTCTTCGATG  GATTCGGCTTAAG  1260
Qy      1275  AATATGCGGATATGGGTTCTCTCGATGCGATGAGCAACACCTC  AGCCAGAACACA  1334
Db      1261  AATATGCGGATATGGGTTCTCTCGATGCGATGAGCAACACCTC  AGCCAGAACACA  1320
Qy      1335  TATTTTCAGTGAAGCTGACAAATCAAAATGCGCCAGAGAGTGTG  CTGTGTCAGGAC  1394
Db      1321  TATTTTCAGTGAAGCTGACAAATCAAAATGCGCCAGAGAGTGTG  CTGTGTCAGGAC  1380
Qy      1395  AAGAGGTCAATCCACAATTTGTCTTACCTGATTGCTGCAATCAA  TACTCATGCCAG  1454
Db      1381  AAGAGGTCAATCCACAATTTGTCTTACCTGATTGCTGCAATCAA  TACTCATGCCAG  1440
Qy      1455  GACATTTGTCAGCAGAGCTTGACCCAGTCCAGGCATGATGACT  TGGGAGCTTAAG  1514
Db      1441  GACATTTGTCAGCAGAGCTTGACCCAGTCCAGGCATGATGACT  TGGGAGCTTAAG  1500
Qy      1515  TTTGAGAGAGAGAGCTCTCAGCCAGGTGGAAGGTGCGTCCATAG  CTCCATTCGTAT  1574
Db      1501  TTTGAGAGAGAGAGCTCTCAGCCAGGTGGAAGGTGCGTCCATAG  CTCCATTCGTAT  1560
Qy      1575  GAGAAGCGCTTTTCTGAAAAGAGGATCCGACACCTCCTCGGTT  TTTTCAATAAAA  1634
Db      1561  GAGAAGCGCTTTTCTGAAAAGAGGATCCGACACCTCCTCGGTT  TTTTCAATAAAA  1620
Qy      1635  GTTTAGAAAAGA  1645
Db      1621  GTTTAGAAAAGA  1631

```

```

RESULT 6
LOCUS    BC015567      1655 bp      mRNA      linear      PRI 09-OCT-2001
DEFINITION Homo sapiens, IMP (inosine monophosphate) dehydrogenase 2, clone
MGC:20947 IMAGE:4576285, mRNA, complete cds.

```

```

ACCESSION BC015567
VERSION   BC015567.1  GI:15990411
KEYWORDS  MGC.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 1655)
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS   Struhsberg R.
JOURNAL   Direct Submission
          Submitted (04-Oct-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Louis Staudt
          CDNA Library Preparation: Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland.
          Web site: http://www.nisc.nih.gov/
          Contact: nisc.mgc@nih.gov
          Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
          Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
          Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,
          Tjongson, E.E., Touchman, J.W., Tsurguev, C., Vogt, J.L., Walker, M.A.,
          Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAL Plate: 30 Row: 3 Column: 8
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis.
FEATURES             Location/Qualifiers
     source            1..1655
     organism="Homo sapiens"
     db_xref="LocusID:3615"
     /db_xref="taxon:9606"
     /clone="MGC:20947 IMAGE:4576285"
     /tissue_type="Primary B-cells from Tonsils"
     /clone_lib="NIH-MGC_48"
     /lab_host="DH10B-R"
     /note="Vector: pORb7"
     41..1585
     CDS               1
     /codon_start=1
     /product="IMP (inosine monophosphate) dehydrogenase 2"
     /protein_id="AAH15567.1"
     /db_xref="GI:15990412"
     /translation="MADYLISGTSYVPDDGLTRAQQLFNCGDLGTYNDPLLTPGYIDF
     TADVDLTSALTLLKTKITLTPVSSPMDDVTAGAMALALGGLGIFIHNNCTPEFOAN
     EVRRKKYKEOGFETDPVLSKRDVDEKARHGPCGIPITPGRGSGPLVGISS
     RDIDFLKEEHDCEFLIEIMTKREDLVVAPACITTLKEANILDRSKKGLPIVNEDEL
     VAITARDLKKRNRYPLASDKAKROILCGAIGTHEDDLYDLIDLAAGVNVVYDSS
     OGNSTFOINMKIKDKYPNIOVIGNNVYTAQAQKNLIDAVDLRPMGSGSCTTIO
     EVLACGRPQATAVTKVSEYARFPGVPVYADGIONVITAFALDLKVAQVSGAVDOK
     TTEAPGEVFEFSDGIRLKRKYRGMGSLDAMDKILSSONRFFSADKIKVAQVSGAVDOK
     GSIHKFEVYLILAGIOHSCODIGAKSLTQVRAMMYSGELKFEKRTSSAOVEGVSHLS
     YEKRLF"
BASE COUNT  430 a      410 c      450 g      365 t
ORIGIN
Query Match      70.7%; Score 1170; DB 9; Length 1655;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1620; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy      17  CGAGACAGCGCGCGGTGCTGTGTCGCAATGGCCGACTGATGATGAGGGGAGC  76
          |||

```

Db	10	CGAGACACGGGGGGGTGTCTCTGTGTGGCCATGGCCCACTACCTGATTATAGTGGGGGAC	69
Oy	77	GTCTTAGTGCCAGAGCAGGACTCACAGCACAGCACTCTTCACTCGGAGAGCGCT	136
Db	70	GTCTCAGCTGGCAAGACGAGACTCACAGCACAGCACTTCACTCGGAGAGCGCT	129
Oy	137	CACCAACAATACCTTTCATTCCTCCCGGTACATCGACTCAGTGCAGACCAAGTGA	196
Db	130	CACCTACATACCTTCTCATCTCTCCCTGGGTACATGACTTACTGACGACCAAGTGA	189
Oy	197	CTGACTCTCTCTGACCAAGAAATACCTCTTAAGACCCCACTGTCTCTCCAT	256
Db	190	CTGACTCTCTCTGACCAAGAAATACCTCTTAAGACCCCACTGTCTCTCCAT	249
Oy	257	GGACACACTCACAAAGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCT	316
Db	250	GGACACACTCACAAAGCTGGGATGGCCATAGCAATGGCGCTTACAGCGGATTTGGCT	309
Oy	317	CATCCACCAACACTGCACCTGAATTCAGGCGCAATGAAGTTGGAAAGTGAACAATA	376
Db	310	CATCCACCAACACTGCACCTGAATTCAGGCGCAATGAAGTTGGAAAGTGAACAATA	369
Oy	377	TGACACGAGATTCAATCAAGACCTGTGTCTCAGCCCCCAAGATCGCGTCGGGATGT	436
Db	370	TGAAACAGGATTCAATCAAGACCTGTGTCTCAGCCCCCAAGATCGCGTCGGGATGT	429
Oy	437	TTTTAGGCCCAAGGCCCGGGCATGGTTTTCTGGGTATCCCAATCACACACAGGCGGAT	496
Db	430	TTTTAGGCCCAAGGCCCGGGCATGGTTTTCTGGGTATCCCAATCACACACAGGCGGAT	489
Oy	497	GGGAGCGCGTGGTGGGCATCATCTCTCAGAGCAATTGATTTTCTCAAGAGAGAGA	556
Db	490	GGGAGCGCGTGGTGGGCATCATCTCTCAGAGCAATTGATTTTCTCAAGAGAGAGA	549
Oy	557	ACATACAGTTCCTTGGAAAGATATGACAAGAGAGGAAGCTGTGCTGACCCCGCC	616
Db	550	ACATACAGTTCCTTGGAAAGATATGACAAGAGAGGAAGCTGTGCTGACCCCGCC	609
Oy	617	CAGCATCACTGGAAGAGGCAATGAAATTCGACAGCGACGACAAGAGAAATTTGCC	676
Db	610	AGGCATCACTGGAAGAGGCAATGAAATTCGACAGCGACGACAAGAGAAATTTGCC	669
Oy	677	CATTGTAAATGAATGATGATGAGCTTGTGGCCATCATTTGCCCGACAGACTTAAGAAGAA	736
Db	670	CATTGTAAATGAATGATGATGAGCTTGTGGCCATCATTTGCCCGACAGACTTAAGAAGAA	729
Oy	737	TGCGGACTACCCACTAGCTCTCCAAAGATGGCCAGAAGAACAGCTGTGTGGGCGAGCAT	796
Db	730	TGCGGACTACCCACTAGCTCTCCAAAGATGGCCAGAAGAACAGCTGTGTGTGGGCGAGCAT	789
Oy	797	TGGCACTCATGAGGATGACAAGATATAGGCTGTGACTTGTCCGCCAGCGTGTGTGATGT	856
Db	790	TGGCACTCATGAGGATGACAAGATATAGGCTGTGACTTGTCCGCCAGCGTGTGTGATGT	849
Oy	857	AGTGGTTTGGACCTTCCCGAGGAAATTCATCTTCGAGATCAATATGATCAACTACAT	916
Db	850	AGTGGTTTGGACCTTCCCGAGGAAATTCATCTTCCGATCAATATGATCAAACTACAT	909
Oy	917	CAAAAGCAAAATACCTTAATCTCCAAGTCATTTGGAGGCAATGTGTGCTAGCTGCCAGGC	976
Db	910	CAAAAGCAAAATACCTTAATCTCCAAGTCATTTGGAGGCAATGTGTGCTAGCTGCCAGGC	969
Oy	977	CAAGACCTCATTTGATATCAGGTGTGGAATGGCCCTGGCGGTGGGATGTGGAAAGTGGCTCAT	1036
Db	970	CAAGACCTCATTTGATATCAGGTGTGGAATGGCCCTGGCGGTGGGATGTGGAAAGTGGCTCAT	1029
Oy	1037	CTGCATTTTTCAGAGAAATGTGTGGCGTGTGGGCGGCCCCCAAGCAACAGCACTTACAAGAT	1096
Db	1030	CTGCATTTTTCAGAGAAATGTGTGGCGTGTGGGCGGCCCCCAAGCAACAGCACTTACAAGAT	1089
Oy	1097	GTAATGATATCAGCGCGCTTGTGGTGTCCGGCTCATTTGCTGATGAGAGAAATCCAATATGT	1156
Db	1090	GTCAGAGATATCAGCGCGCTTGTGGTGTCCGGCTCATTTGCTGATGAGAGAAATCCAATATGT	1149

Qy	1157	GGGTCAATATTTGGGAAAGCCTTTGGCCCTTGGGGCCCTCCACAGTCATGATGGGCTCTCTCT	1216
Db	1150	GGGTCAATATTTGGGAAAGCCTTTGGCCCTTGGGGCCCTCCACAGTCATGATGGGCTCTCTCT	1209
Qy	1217	GGCTCCACACCATGAGGGCCCGTGGTAATCTCTTTTCCGATGGGATCCGGCTAAAGAA	1276
Db	1210	GGCTCCACACCATGAGGGCCCGTGGTAATCTCTTTTCCGATGGGATCCGGCTAAAGAA	1269
Qy	1277	ATATGCGGATAGGGTTCTCTCGATGCCATGAGCAAGCACCTCAGCACAGCAGAAACAGATA	1336
Db	1270	ATATGCGGATAGGGTTCTCTCGATGCCATGAGCAAGCACCTCAGCACAGCAGAAACAGATA	1329
Qy	1337	TTTTCAGTGAAGTGCACAAATATCAAGTGCGCCAGGGAATGCTGTGCTGTGTCAGAGCAA	1396
Db	1330	TTTTCAGTGAAGTGCACAAATATCAAGTGCGCCAGGGAATGCTGTGCTGTGTCAGAGCAA	1389
Qy	1397	ACGGTCATTCACCAATTTTGCTTACCTGATTCCTGTCGATCCCAACACTCATGCGCAGGA	1456
Db	1390	ACGGTCATTCACCAATTTTGCTTACCTGATTCCTGTCGATCCCAACACTCATGCGCAGGA	1449
Qy	1457	CATTGGTGCCCAAGAGCTTGACCCCAAGTCCGAGCCATGATGTACTCTGGGGAGCTTAAATT	1516
Db	1450	CATTGGTGCCCAAGAGCTTGACCCCAAGTCCGAGCCATGATGTACTCTGGGGAGCTTAAATT	1509
Qy	1517	TTGAGAAGAGACGCTCTCAGCCAGGTGGAAGTGGCTCATAGCTCCATTCGTATGA	1576
Db	1510	TTGAGAAGAGAGCTCTCAGCCAGGTGGAAGTGGCTCATAGCTCCATTCGTATGA	1569
Qy	1577	GAAGCGCTTTTCTGAAAGGATGCACACACCTCTCGGTTTTTTTTTTCATATAAAGT	1636
Db	1570	GAAGCGCTTTTCTGAAAGGATGCACACACCTCTCGGTTTTTTTTTTCATATAAAGT	1629
Qy	1637	TTTGAAGAAG 1645	
Db	1630	TTTGAAGAAG 1638	
RESULT 7			
LOCUS	AX418035	1158 bp	Linear
DEFINITION	Sequence 45 from Patent WO0185952.	DNA	PAT 18-JUN-2002
ACCESSION	AX418035		
VERSION	AX418035.1	GI:21523082	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 Krystek,S.R., Sherliff,S., Wilmer,M.R., Hollenbaugh,D.L., Yan,N., Mouravieff,J.E., Elmspahr,H.M. and Kish,K.		
AUTHORS	Modified inosine 5'-monophosphate dehydrogenase polypeptides and uses thereof		
TITLE	Patent: WO 0185952-A 45 15-NOV-2001;		
JOURNAL	Bristol-Myers Squibb Co. (US)		
FEATURES	Location/Qualifiers		
source	1..1158 /organism="Homo sapiens" /db_xref="taxon:9606"		
BASE COUNT	279 a	301 c	310 g 268 t
ORIGIN			
Query Match	37.2%	Score 615;	DB 6; Length 1158;
Best Local Similarity	99.5%;	Pred. No. 0;	
Matches 815;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	774	CAGCGTGTGTGGGGGACCATTTGGCAGCTCATGAGATGACAGTATAGGCTGGACTTG	833
Db	340	CAGCTGTGTGTGGGGGACCATTTGGCAGCTCATGAGATGACAGTATAGGCTGGACTTG	399
Qy	834	CTCCGCCAGGCGTGTGTGATGTAGTGGTTTGGACTTTCCACAGGAAATTCCATCTTC	893

Db 400 CTCGCCACAGCTGCTGTGATGATGAGTGTGTTGGACTCTTCCAGAGGAATTCCATCTTC 459  
Qy 894 CAGATCAATATGATCAAGTACATCAAGACAATAACCTAATCTCCAAATCATTTGGAGGC 953  
Db 460 CAGATCAATATGATCAAGTACATCAAGACAATAACCTAATCTCCAAATCATTTGGAGGC 519  
Qy 954 AATGAGTACAGTGTGCCAGGCCAAGACCTATTGATGACAGGTGTGATGCCCTGGCG 1013  
Db 520 AATGAGTACAGTGTGCCAGGCCAAGACCTATTGATGACAGGTGTGATGCCCTGGCG 579  
Qy 1014 GTGGGCATGGAAGTGGCTTCATCTGATTTATCCAGGAAGTCTGTGCC 1073  
Db 580 GTGGGCATGGAAGTGGCTTCATCTGATTTATCCAGGAAGTCTGTGCC 639  
Qy 1074 CAGGCAACGAGTGTACAGAGGTATGATGACAGGCGCTTGGTGTTCGGGCTATT 1133  
Db 640 CAGGCAACGAGTGTACAGAGGTATGATGACAGGCGCTTGGTGTTCGGGCTATT 699  
Qy 1134 CCTGATGGAAGATCCAAATGTGGGTCAATTATGCGAAAGCCTTGGCC 1193  
Db 700 CCTGATGGAAGATCCAAATGTGGGTCAATTATGCGAAAGCCTTGGCC 759  
Qy 1194 ACAGTCATGATGGCTCTCTCTGCTGCCACCACTGAGGCGCCCTGGT 1253  
Db 760 ACAGTCATGATGGCTCTCTCTGCTGCCACCACTGAGGCGCCCTGGT 819  
Qy 1254 TCCGATGGGATCCGGCTAAGAATATCCGGTATGGTCTCTCGAT 1313  
Db 820 TCCGATGGGATCCGGCTAAGAATATCCGGTATGGTCTCTCGAT 879  
Qy 1314 CACCTCAGCAGCAGACAAGATATTTCAGTGAAGCTGACAATAATCAA 1373  
Db 880 CACCTCAGCAGCAGACAAGATATTTCAGTGAAGCTGACAATAATCAA 939  
Qy 1374 GTGTCTGTGCTGTGCAGACAAAAGGTCATCCAAATTTGTCTCCT 1433  
Db 940 GTGTCTGTGCTGTGCAGACAAAAGGTCATCCAAATTTGTCTCCT 999  
Qy 1434 GGCATCAACACTCATTCGAGGACATTGGTGCAGAGCTTGACCCAA 1493  
Db 1000 GGCATCAACACTCATTCGAGGACATTGGTGCAGAGCTTGACCCAA 1059  
Qy 1494 AATGACTCTGGGAGCTTAAGTTTGAGAAGAACGCTCTCAGCCAG 1553  
Db 1060 AATGACTCTGGGAGCTTAAGTTTGAGAAGAACGCTCTCAGCCAG 1119  
Qy 1554 GTCCATAGCCTCATTCGTATGAGAAGCGGCTTTCTGA 1592  
Db 1120 GTCCATAGCCTCATTCGTATGAGAAGCGGCTTTCTGA 1158

RESULT 8  
AX418036 1158 bp DNA linear PAT 18-JUN-2002  
LOCUS Sequence 46 from Patent W00185952.  
DEFINITION AX418036  
ACCESSION AX418036.1 GI:21523083  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1  
AUTHORS Krystek,S.R., Sherliff,S., Witmer,M.R., Hollenbaugh,D.L., Yan,N.,  
Moudaviet,J.E., Einspahr,H.M. and Kish,K.  
TITLE Modified inosine 5'-monophosphate dehydrogenas polypeptides and  
uses thereof  
JOURNAL Patent: WO 0185952-A 46 15-NOV-2001;  
Bristol-Myers Squibb Co. (US)  
FEATURES  
SOURCE Location/Qualifiers  
1. 1158  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 277 a 300 c 313 g 268 t  
ORIGIN  
Query Match 37.1%; Score 613; DB 6; Length 1158;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 813; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 776 GCTGCTGTGGGGGAGCCATTGGCACTCATAGATGACAGTATAGCGTGGACTTGT 835  
Db 342 GCTGCTGTGGGGAGCCATTGGCACTCATAGATGACAGTATAGCGTGGACTTGT 401  
Qy 836 CCGCCAGCGCTGTGTGATGTAGTGTGTTTGACTCTTCCAGGGAATTTCCATCTTCA 895  
Db 402 CCGCCAGCGCTGTGTGATGTAGTGTGTTTGACTCTTCCAGGGAATTTCCATCTTCA 461  
Qy 896 GATCAATATGATCAAGTACATCAAGACAATACCTAATCTCCAAAGTATTGGAGCAA 955  
Db 462 GATCAATATGATCAAGTACATCAAGACAATACCTAATCTCCAAAGTATTGGAGCAA 521  
Qy 956 TGTGCTACTGTGCCAGGCAAGAACCTCATTTGATGAGGTGTGGATGCCCTGGGGGT 1015  
Db 522 TGTGCTACTGTGCCAGGCAAGAACCTCATTTGATGAGGTGTGGATGCCCTGGGGGT 581  
Qy 1016 GGGCATGGGAAGTGGCTCCATCTGCTTATCCAGCAAGTGTGGCTGTGGGCGCCCA 1075  
Db 582 GGGCATGGGAAGTGGCTCCATCTGCTTATCCAGCAAGTGTGGCTGTGGGCGCCCA 641  
Qy 1076 AGCAACAGCAGTGTACAAGTGTATGATGATGACAGCGGCTTGGTGTCCGTCATTGC 1135  
Db 642 AGCAACAGCAGTGTACAAGTGTATGATGATGACAGCGGCTTGGTGTCCGTCATTGC 701  
Qy 1136 TGATGAGGAATTCAAATATGTGGGTCAATTGGGAAAGCCTTGGCCCTGGGCGCTCCAC 1195  
Db 702 TGATGAGGAATTCAAATATGTGGGTCAATTGGGAAAGCCTTGGCCCTGGGCGCTCCAC 761  
Qy 1196 AGTCATGATGGGCTCTCTCTGCTGCTGACCACTGAGGCGCCCTGGTGAATTTCTTTTC 1255  
Db 762 AGTCATGATGGGCTCTCTCTGCTGCTGACCACTGAGGCGCCCTGGTGAATTTCTTTTC 821  
Qy 1256 CGATGGGATCCGGCTTAAGAAATATCCGGGTATGGGCTTCTCGATGCCATGACAGCA 1315  
Db 822 CGATGGGATCCGGCTTAAGAAATATCCGGGTATGGGCTTCTCGATGCCATGACAGCA 881  
Qy 1316 CCTCAGCAGCCAGAACAGATATTTCAGTGAAGCTGACAATAATCAAGTGCCAGAGAGT 1375  
Db 882 CCTCAGCAGCCAGAACAGATATTTCAGTGAAGCTGACAATAATCAAGTGCCAGAGAGT 941  
Qy 1376 GTGTGCTGTGTGTCAGAGCAAAAGGTCATTCACAAATTTGTCCCTTACTGATTTGCTGG 1435  
Db 942 GTGTGCTGTGTGTCAGAGCAAAAGGTCATTCACAAATTTGTCCCTTACTGATTTGCTGG 1001  
Qy 1436 CATCAACACTCATTCGAGGACATTGGTGCAGGACCTTGACCCAGGACCTGAT 1495  
Db 1002 CATCAACACTCATTCGAGGACATTGGTGCAGGACCTTGACCCAAAGTGTCCGAGCATTGAT 1061  
Qy 1496 GTACTCTGGGAGCTTAAGTTTGAGAAGAACGTCCTACGCCAGGTGGAAGGTGGCGT 1555  
Db 1062 GTACTCTGGGAGCTTAAGTTTGAGAAGAACGTCCTACGCCAGGTGGAAGGTGGCGT 1121  
Qy 1556 CCATAGCCTTCATTCGTATGAGAAGCGGCTTTCTGA 1592  
Db 1122 CCATAGCCTTCATTCGTATGAGAAGCGGCTTTCTGA 1158

RESULT 9  
AX418030 1155 bp DNA linear PAT 18-JUN-2002  
LOCUS Sequence 40 from Patent W00185952.  
DEFINITION AX418030  
ACCESSION AX418030.1 GI:21523077  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1  
REFERENCE 1  
AUTHORS 1  
TITLE 1  
JOURNAL 1  
FEATURES 1  
SOURCE 1  
BASE COUNT 281 a 299 c 310 g 265 t  
ORIGIN 1  
Query Match 37.0%; Score 612; DB 6; Length 1155;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 777 CTGCTGTGTGGGCGACCCATTGGCACTCATGAGATGACAGATATGAGCTGGACTTGCCTC 836  
DB 340 CTGCTGTGTGGGCGACCCATTGGCACTCATGAGATGACAGATATGAGCTGGACTTGCCTC 399  
QY 837 GCCCAGGCTGGTGTGATAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 896  
DB 400 GCCCAGGCTGGTGTGATAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 459  
QY 897 ATCAATATGATCAAGTACATCAAAAGCAATACCTTAATCTCCAAGTCAATTGGAGCAAT 956  
DB 460 ATCAATATGATCAAGTACATCAAAAGCAATACCTTAATCTCCAAGTCAATTGGAGCAAT 519  
QY 957 GTGTGACTGTGCCAGGCCAAGAACTCATTTGATGACAGTGTGATGCTCCCTGGGGGTG 1016  
DB 520 GTGTGACTGTGCCAGGCCAAGAACTCATTTGATGACAGTGTGATGCTCCCTGGGGGTG 579  
QY 1017 GGCATGGAAGTGGCTCCATCTGCATTATTCAGGAAGTCTGGGCTGTGGGGGCCCAA 1076  
DB 580 GGCATGGAAGTGGCTCCATCTGCATTATTCAGGAAGTCTGGGCTGTGGGGGCCCAA 639  
QY 1077 GCAACAGCAGTGTACAAAGTGTATGATGACAGGCGCTTGGTGTCCGCTCATTTGCT 1136  
DB 640 GCAACAGCAGTGTACAAAGTGTATGATGACAGGCGCTTGGTGTCCGCTCATTTGCT 699  
QY 1137 GATGAGAGCAATCAAAATTTGGGTCATATTTGCAAAAGCTTGGGCTGGGCTCCACA 1196  
DB 700 GATGAGAGCAATCAAAATTTGGGTCATATTTGCAAAAGCTTGGGCTGGGCTCCACA 759  
QY 1197 GTCATGATGGGCTCTCTCTGCTGCTGCACACTGAGGCGCTGTGTGATATCTTTTCC 1256  
DB 760 GTCATGATGGGCTCTCTCTGCTGCTGCACACTGAGGCGCTGTGTGATATCTTTTCC 819  
QY 1257 GATGGGATCGGCTTAAAGAAATATGCGGCTATGGGTCCTGATGGCCATGCAACAGC 1316  
DB 820 GATGGGATCGGCTTAAAGAAATATGCGGCTATGGGTCCTGATGGCCATGCAACAGC 879  
QY 1317 CTCACAGCAGCAGAACAGATATTTCACTGAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTG 1376  
DB 880 CTCACAGCAGCAGAACAGATATTTCACTGAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTG 939  
QY 1377 TCTGTGCTGTGACAGACAAAGGTCATCACAAAATTTGTCCTTACCTGATTTGCTGSC 1436  
DB 940 TCTGTGCTGTGACAGACAAAGGTCATCACAAAATTTGTCCTTACCTGATTTGCTGSC 999  
QY 1437 ATCCAAACCTCATTTGCGAGACATTTGGTGGCAGAGCTTACCCCAAGTCCGAGCCATGATG 1496  
DB 1000 ATCCAAACCTCATTTGCGAGACATTTGGTGGCAGAGCTTACCCCAAGTCCGAGCCATGATG 1059  
QY 1497 TACTGTGGGAGCTTAAGTTTGAAGAAGAACGTCCTCAGCCAGGTGAAGTGGAGCTG 1556  
DB 1060 TACTGTGGGAGCTTAAGTTTGAAGAAGAACGTCCTCAGCCAGGTGAAGTGGAGCTG 1119  
QY 1557 CATAGCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1592

Db 1120 CATAGCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1155  
RESULT 10  
LOCUS AX418031 1155 bp DNA Linear PAT 18-JUN-2002  
DEFINITION Sequence 41 from Patent WO0185952.  
ACCESSION AX418031  
VERSION AX418031.1 GI:21523078  
KEYWORDS  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1  
REFERENCE 1  
AUTHORS 1  
TITLE 1  
JOURNAL 1  
FEATURES 1  
SOURCE 1  
BASE COUNT 278 a 300 c 310 g 267 t  
ORIGIN 1  
Query Match 37.0%; Score 612; DB 6; Length 1155;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 777 CTGCTGTGTGGGCGACCCATTGGCACTCATGAGATGACAGATATGAGCTGGACTTGCCTC 836  
DB 340 CTGCTGTGTGGGCGACCCATTGGCACTCATGAGATGACAGATATGAGCTGGACTTGCCTC 399  
QY 837 GCCCAGGCTGGTGTGATAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 896  
DB 400 GCCCAGGCTGGTGTGATAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 459  
QY 897 ATCAATATGATCAAGTACATCAAAAGCAATACCTTAATCTCCAAGTCAATTGGAGCAAT 956  
DB 460 ATCAATATGATCAAGTACATCAAAAGCAATACCTTAATCTCCAAGTCAATTGGAGCAAT 519  
QY 957 GTGTGACTGTGCCAGGCCAAGAACTCATTTGATGACAGTGTGATGCTCCCTGGGGGTG 1016  
DB 520 GTGTGACTGTGCCAGGCCAAGAACTCATTTGATGACAGTGTGATGCTCCCTGGGGGTG 579  
QY 1017 GGCATGGAAGTGGCTCCATCTGCATTATTCAGGAAGTCTGGGCTGTGGGGGCCCAA 1076  
DB 580 GGCATGGAAGTGGCTCCATCTGCATTATTCAGGAAGTCTGGGCTGTGGGGGCCCAA 639  
QY 1077 GCAACAGCAGTGTACAAAGTGTATGATGACAGGCGCTTGGTGTCCGCTCATTTGCT 1136  
DB 640 GCAACAGCAGTGTACAAAGTGTATGATGACAGGCGCTTGGTGTCCGCTCATTTGCT 699  
QY 1137 GATGAGAGCAATCAAAATTTGGGTCATATTTGCAAAAGCTTGGGCTGGGCTCCACA 1196  
DB 700 GATGAGAGCAATCAAAATTTGGGTCATATTTGCAAAAGCTTGGGCTGGGCTCCACA 759  
QY 1197 GTCATGATGGGCTCTCTCTGCTGCTGCACACTGAGGCGCTGTGTGATATCTTTTCC 1256  
DB 760 GTCATGATGGGCTCTCTCTGCTGCTGCACACTGAGGCGCTGTGTGATATCTTTTCC 819  
QY 1257 GATGGGATCGGCTTAAAGAAATATGCGGCTATGGGTCCTGATGGCCATGCAACAGC 1316  
DB 820 GATGGGATCGGCTTAAAGAAATATGCGGCTATGGGTCCTGATGGCCATGCAACAGC 879  
QY 1317 CTCACAGCAGCAGAACAGATATTTCACTGAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTG 1376  
DB 880 CTCACAGCAGCAGAACAGATATTTCACTGAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTG 939

```
QY 1377 TCTGTCGCTGTGCAGCAAAAGGTCATCCAAATTTGCTCCCTAGCTGATTTGCTGGC 1436
DB 940 TCTGTCGCTGTGCAGCAAAAGGTCATCCAAATTTGCTCCCTAGCTGATTTGCTGGC 999
QY 1437 ATCCAACTCATGATGACGACATTTGGTCCCAAGAGCTTGACCCCAAGTCCAGCCATGATG 1496
DB 1000 ATCCAACTCATGATGACGACATTTGGTCCCAAGAGCTTGACCCCAAGTCCAGCCATGATG 1059
QY 1497 TACTCTGGGAGGCTTAAGTTTGAGAGAGAGAGCTCCTCAGCCAGGCTGAAGTGGCGTC 1556
DB 1060 TACTCTGGGAGGCTTAAGTTTGAGAGAGAGAGCTCCTCAGCCAGGCTGAAGTGGCGTC 1119
QY 1557 CATAGCCTCCATTGATGATGAGAACGCGCTTTCTGA 1592
DB 1120 CATAGCCTCCATTGATGATGAGAACGCGCTTTCTGA 1155

RESULT 11
AX418032
LOCUS AX418032 1155 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 42 from Patent WO0185952.
ACCESSION AX418032
VERSION AX418032.1 GI:21523079
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 Krystek,S.R., Sheriff,S., Witmer,M.R., Hollenbaugh,D.L., Yan,N.,
Mouravieff,J.E., Einspahr,H.M. and Kish,K.
Modified inosine 5'-monophosphate dehydrogenase polypeptides and
uses thereof
Patent: WO 0185952-A 42 15-NOV-2001:
Bristol-Myers Squibb Co. (US)
LOCATION/Qualifiers
1. 1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 277 a 299 c 312 g 267 t
ORIGIN
Query Match 37.0%; Score 612; DB 6; Length 1155;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 777 CTGCTGTGGGGGACGACATTTGGCACTCATGAGATGACAGATATAGGCTGGACTTGCCTC 836
DB 340 CTGCTGTGGGGGACGACATTTGGCACTCATGAGATGACAGATATAGGCTGGACTTGCCTC 399
QY 837 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATCCATCTTCCAG 896
DB 400 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATCCATCTTCCAG 459
QY 897 ATCAATATGATCAAGTATCAAGACAATATACCTTAATCTCCAACTCTTGGAGCAAT 956
DB 460 ATCAATATGATCAAGTATCAAGACAATATACCTTAATCTCCAACTCTTGGAGCAAT 519
QY 957 GTGTCACCTGCTGCCAGGCGCAAGAACCTCATTTGATGACAGTGTGATGCCCTGCGGCGGTG 1016
DB 520 GTGTCACCTGCTGCCAGGCGCAAGAACCTCATTTGATGACAGTGTGATGCCCTGCGGCGGTG 579
QY 1017 GGCATGGGAAGTGGTCTCATCTTATTCACAGAAAGTGTGCTGGCGGCCCA 1076
DB 580 GGCATGGGAAGTGGTCTCATCTTATTCACAGAAAGTGTGCTGGCGGCCCA 639
QY 1077 GCAACAGCAGTGTACAGGTGTATGATATGACGAGCGGCTTTGGTGTTCGGTCATTGGCT 1136
DB 640 GCAACAGCAGTGTACAGGTGTATGATATGACGAGCGGCTTTGGTGTTCGGTCATTGGCT 699
QY 1137 GATGGAGGAATCCAAATGTGGTCAATTTGGCAAGCCTTGGCCCTTGGGCGCTCCACA 1196
DB 700 GATGGAGGAATCCAAATGTGGTCAATTTGGCAAGCCTTGGCCCTTGGGCGCTCCACA 759
```

```
QY 1197 GTCATGATGGGCTCTCTCTGCTGCCACCTGAGAGCCCTTGTAATTACTTTTCC 1256
DB 760 GTCATGATGGGCTCTCTCTGCTGCCACCTGAGAGCCCTTGTAATTACTTTTCC 819
QY 1257 GATGGATCCGGCTTAAAGAAATATCGGGATGAGGTTCTCTCATGATGACATGACAGCAC 1316
DB 820 GATGGATCCGGCTTAAAGAAATATCGGGATGAGGTTCTCTCATGATGACATGACAGCAC 879
QY 1317 CTCAGCAGCCAGAACAGATATTTTCAGTGAAGCTGACAAATAAAGTGGCCAGGAGTG 1376
DB 880 CTCAGCAGCCAGAACAGATATTTTCAGTGAAGCTGACAAATAAAGTGGCCAGGAGTG 939
QY 1377 TCTGTCGCTGTGCAGCAAAAGGTCATCCAAATTTGCTCCCTAGCTGATTTGCTGGC 1436
DB 940 TCTGTCGCTGTGCAGCAAAAGGTCATCCAAATTTGCTCCCTAGCTGATTTGCTGGC 999
QY 1437 ATCCAACTCATGATGACGACATTTGGTCCCAAGAGCTTGACCCCAAGTCCAGCCATGATG 1496
DB 1000 ATCCAACTCATGATGACGACATTTGGTCCCAAGAGCTTGACCCCAAGTCCAGCCATGATG 1059
QY 1497 TACTCTGGGAGGCTTAAGTTTGAGAGAGAGAGCTCCTCAGCCAGGCTGAAGTGGCGTC 1556
DB 1060 TACTCTGGGAGGCTTAAGTTTGAGAGAGAGAGCTCCTCAGCCAGGCTGAAGTGGCGTC 1119
QY 1557 CATAGCCTCCATTGATGATGAGAACGCGCTTTCTGA 1592
DB 1120 CATAGCCTCCATTGATGATGAGAACGCGCTTTCTGA 1155

RESULT 12
AX418033
LOCUS AX418033 1155 bp DNA linear PAT 19-JUN-2002
DEFINITION Sequence 43 from Patent WO0185952.
ACCESSION AX418033
VERSION AX418033.1 GI:21523080
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 Krystek,S.R., Sheriff,S., Witmer,M.R., Hollenbaugh,D.L., Yan,N.,
Mouravieff,J.E., Einspahr,H.M. and Kish,K.
Modified inosine 5'-monophosphate dehydrogenase polypeptides and
uses thereof
Patent: WO 0185952-A 43 15-NOV-2001:
Bristol-Myers Squibb Co. (US)
LOCATION/Qualifiers
1. 1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 278 a 300 c 309 g 268 t
ORIGIN
Query Match 37.0%; Score 612; DB 6; Length 1155;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 777 CTGCTGTGGGGGACGACATTTGGCACTCATGAGATGACAGATATAGGCTGGACTTGCCTC 836
DB 340 CTGCTGTGGGGGACGACATTTGGCACTCATGAGATGACAGATATAGGCTGGACTTGCCTC 399
QY 837 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 896
DB 400 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCCAG 459
QY 897 ATCAATATGATCAAGTATCAAGACAATATACCTTAATCTCCAACTCTTGGAGCAAT 956
DB 460 ATCAATATGATCAAGTATCAAGACAATATACCTTAATCTCCAACTCTTGGAGCAAT 519
QY 957 GTGTCACCTGCTGCCAGGCGCAAGAACCTCATTTGATGACAGTGTGATGCCCTTGGGCGGTG 1016
DB 700 GTGTCACCTGCTGCCAGGCGCAAGAACCTCATTTGATGACAGTGTGATGCCCTTGGGCGGTG 1016
```

Db 520 GTGTCACCTGCTGCCCCAGAGCCATTTGATGCAGGTGTGTGAGATGCCCTGCGGGTG 579  
Oy 1017 GGCATGGGAAGTGGCTCCATCGATTTATCCAGGAAGTCTGGCTGTGGGCGGCCCAA 1076  
Db 580 GGCATGGGAAGTGGCTCCATCGATTTAGCGAGGAAGTCTGGCTGTGGGCGGCCCAA 639  
Oy 1077 GCACAGCAGTGTACAAAGTGTATGAGTATGCAGCGCGCTTTGTGTCCGGTCATTGCT 1136  
Db 640 GCACAGCAGTGTACAAAGTGTATGAGTATGCAGCGCGCTTTGTGTCCGGTCATTGCT 699  
Oy 1137 GATGAGGAATCCAAATGTGGGTCATATTGGAAAGCTTGGCCCTGGGGCCCTCCACA 1196  
Db 700 GATGAGGAATCCAAATGTGGGTCATATTGCCAAAGCTTGGCCCTGGGGCCCTCCACA 759  
Oy 1197 GTCATGATGGGCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATACTTCTTTTCC 1256  
Db 760 GTCATGATGGGCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATACTTCTTTTCC 819  
Oy 1257 GATGGGATCCGGCTAAAGAAATATCGCGGTATGGCTTCTCTGATGCCATGACAAAC 1316  
Db 820 GATGGGATCCGGCTAAAGAAATATCGCGGTATGGCTTCTCTGATGCCATGACAAAC 879  
Oy 1317 CTCACAGCCAGCAACAGATATTTCAGTGAAGCTGACAAATCAAGTGGCCGAGGGAGTG 1376  
Db 880 CTCACAGCCAGCAACAGATATTTCAGTGAAGCTGACAAATCAAGTGGCCGAGGGAGTG 939  
Oy 1377 TCTGGTGTGTGTCAGAGCAAAAGGTCATCCAAATTTGTCCCTTACTGATTTGCTGCGC 1436  
Db 940 TCTGGTGTGTGTCAGAGCAAAAGGTCATCCAAATTTGTCCCTTACTGATTTGCTGCGC 999  
Oy 1437 ATCCAACTCATATGCCAGACATTTGTCGCAAGAGCTTACCCAGTCCGAGCCATGATG 1496  
Db 1000 ATCCAACTCATATGCCAGACATTTGTCGCAAGAGCTTACCCAGTCCGAGCCATGATG 1059  
Oy 1497 TACTCTGGGAGCTTAAGTTTGAGAGAGACGCTCCTCAGCCAGGTGGAAGTGGCGTC 1556  
Db 1060 TACTCTGGGAGCTTAAGTTTGAGAGAGACGCTCCTCAGCCAGGTGGAAGTGGCGTC 1119  
Oy 1557 CATAGCCTCATTCGTATGAGAAGCGGCTTTTCTGA 1592  
Db 1120 CATAGCCTCATTCGTATGAGAAGCGGCTTTTCTGA 1155

RESULT 13  
AX418037 1158 bp DNA linear PAT 18-JUN-2002  
LOCUS Sequence 47 from Patent W00185952.  
DEFINITION AX418037  
ACCESSION AX418037  
VERSION AX418037.1 GI:21523084  
KEYWORDS  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1  
AUTHORS Krystek,S.R., Sheriff,S., Witmer,M.R., Hollenbaugh,D.L., Yan,N.,  
Mounavieff,J.E., Einspahr,H.M. and Kish,K.  
TITLE Modified inosine 5'-monophosphate dehydrogenase polypeptides and  
uses thereof  
JOURNAL Patent: WO 0185952-A 47 15-NOV-2001;  
Bristol-Myers Squibb Co (US)  
FEATURES  
source  
1. 1158  
Location/Qualifiers  
BASE COUNT 279 a 301 c 309 g 269 t  
ORIGIN  
Query Match 37.0%; Score 612; DB 6; Length 1158;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 777 CTGCTGTGTGGGCGACCATTTGGCACTCATGAGATGACAAATGATAGCTGCACTTGCTC 836

Db 343 CTGCTGTGTGGGCGACCATTTGGCACTCATGAGATGACAAATGATAGCTGCACTTGCTC 402  
Oy 837 GCCAGGCTGTGTGGATGTAGTGGTTTGGACTCTTCCAGAGAAATTCATCTTTCAG 896  
Db 403 GCCAGGCTGTGTGGATGTAGTGGTTTGGACTCTTCCAGAGAAATTCATCTTTCAG 462  
Oy 897 ATCAATATGATCACTACATCAACAACAATACCTTAATCTTCAAGTCAATTGGAGCAAT 956  
Db 463 ATCAATATGATCACTACATCAACAACAATACCTTAATCTTCAAGTCAATTGGAGCAAT 522  
Oy 957 GTGTGACCTGTGCCAGGCCCAAGAACCTCATTTGATGACAGGTGTGATGCCCTGGGGTG 1016  
Db 523 GTGTGACCTGTGCCAGGCCCAAGAACCTCATTTGATGACAGGTGTGATGCCCTGGGGTG 582  
Oy 1017 GGCATGGGAAGTGGCTCCATCGATTTAGCGAGGAAGTCTGGCTGTGGGCGGCCCAA 1076  
Db 583 GGCATGGGAAGTGGCTCCATCGATTTAGCGAGGAAGTCTGGCTGTGGGCGGCCCAA 642  
Oy 1077 GCACAGCAGTGTACAAAGTGTATGAGTATGCACGGCGCTTTGTGTCCGGTCATTGCT 1136  
Db 643 GCACAGCAGTGTACAAAGTGTATGAGTATGCACGGCGCTTTGTGTCCGGTCATTGCT 702  
Oy 1137 GATGAGGAATCCAAATGTGGGTCATATTGGAAAGCTTGGCCCTGGGGCCCTCCACA 1196  
Db 703 GATGAGGAATCCAAATGTGGGTCATATTGGAAAGCTTGGCCCTGGGGCCCTCCACA 762  
Oy 1197 GTCATGATGGGCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATACTTCTTTTCC 1256  
Db 763 GTCATGATGGGCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATACTTCTTTTCC 822  
Oy 1257 GATGGATCCGGCTAAAGAAATATCGCGGTATGGCTTCTCTGATGCCATGACAAAC 1316  
Db 823 GATGGATCCGGCTAAAGAAATATCGCGGTATGGCTTCTCTGATGCCATGACAAAC 882  
Oy 1317 CTCACAGCCAGCAACAGATATTTCAGTGAAGCTGACAAATCAAGTGGCCGAGGGAGTG 1376  
Db 883 CTCACAGCCAGCAACAGATATTTCAGTGAAGCTGACAAATCAAGTGGCCGAGGGAGTG 942  
Oy 1377 TCTGGTGTGTGTCAGAGCAAAAGGTCATCCAAATTTGTCCCTTACTGATTTGCTGCGC 1436  
Db 943 TCTGGTGTGTGTCAGAGCAAAAGGTCATCCAAATTTGTCCCTTACTGATTTGCTGCGC 1002  
Oy 1437 ATCCAACTCATATGCCAGACATTTGTCGCAAGAGCTTACCCAGTCCGAGCCATGATG 1496  
Db 1003 ATCCAACTCATATGCCAGACATTTGTCGCAAGAGCTTACCCAGTCCGAGCCATGATG 1062  
Oy 1497 TACTCTGGGAGCTTAAGTTTGAGAGAGACGCTCCTCAGCCAGGTGGAAGTGGCGTC 1556  
Db 1063 TACTCTGGGAGCTTAAGTTTGAGAGAGACGCTCCTCAGCCAGGTGGAAGTGGCGTC 1122  
Oy 1557 CATAGCCTCATTCGTATGAGAAGCGGCTTTTCTGA 1592  
Db 1123 CATAGCCTCATTCGTATGAGAAGCGGCTTTTCTGA 1158

RESULT 14  
AX093346 396 bp DNA linear PAT 30-MAR-2001  
LOCUS AX093346  
DEFINITION Sequence 164 from Patent W00118046.  
ACCESSION AX093346  
VERSION AX093346.1 GI:13509795  
KEYWORDS  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 396)  
AUTHORS Xu,J. and Stolk,J.A.  
TITLE Ovarian tumor sequences and methods of use therefor  
JOURNAL Patent: WO 0118046-A 164 15-MAR-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers

```
source 1. .396
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 96 a 119 c 96 g 85 t
ORIGIN
Query Match 23.9%; Score 396; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 7e-223;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 GACACCGGGGGTGTCTCTGTGTGGCCATGCGGACCTGATTAAGTATGCGGCGACGTC 80
    |||
DB 1 GACACCGGGGGTGTCTCTGTGTGGCCATGCGGACCTGATTAAGTATGCGGCGACGTC 60
    |||
OY 81 TAGTGCCGAGACGAGACGACGACGACGACCTCTTCAACTGGCGGATGCGGCGCTCACC 140
    |||
DB 61 TAGTGCCGAGACGAGACGACGACGACGACCTCTTCAACTGGCGGATGCGGCGCTCACC 120
    |||
OY 141 TACAATGACTTTCCTCTCTCTGCTGGGTACATGACTTCACTGACAGAGAGCTGACCTG 200
    |||
DB 121 TACAATGACTTTCCTCTCTCTGCTGGGTACATGACTTCACTGACAGAGAGCTGACCTG 180
    |||
OY 201 ACTTGTGCTGTGACCAAGAAATCATTAAAGCCCACTGGTTCTCTCCCATGGAC 260
    |||
DB 181 ACTTGTGCTGTGACCAAGAAATCATTAAAGCCCACTGGTTCTCTCCCATGGAC 240
    |||
OY 261 ACAGTCACAGAGGCTGGATGGCCATAGCAANTGGCGCTTACAGGCGCTATTGGCTTCATC 320
    |||
DB 241 ACAGTCACAGAGGCTGGATGGCCATAGCAANTGGCGCTTATTGGCTTCATC 300
    |||
OY 321 CACCACACGTCGTACACCTGATTCGAGGCCAATGAAGTTGGGAAGTGAAGAATATGAA 380
    |||
DB 301 CACCACACGTCGTACACCTGATTCGAGGCCAATGAAGTTGGGAAGTGAAGAATATGAA 360
    |||
OY 381 CAGGATTCATCACAGACCCCTGTGTCTCAGCCCC 416
    |||
DB 361 CAGGATTCATCACAGACCCCTGTGTCTCAGCCCC 396
    |||

RESULT 15
LOCUS AX409550 6193 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2197 from Patent WO0229103.
ACCESSION AX409550
VERSION AX409550.1 GI:21442255
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Alvarres,C., Horne,D., Peres-da-Silva,S. and Vokhley,J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2197 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
source 1. .6193
/organism="Homo sapiens"
/db_xref="taxon:9606"
/locus="EMBL/GenBank Accession No. U33412"
BASE COUNT 1404 a 1597 c 1628 g 1564 t
ORIGIN
Query Match 14.6%; Score 242; DB 6; Length 6193;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1101 GAGTATGCAACCGGCTTTGGTGTCCGTCATTCGATGAGAGAGATTC AAAATGTGGGT 1160
    |||
DB 5210 GAGTATGCAACCGGCTTTGGTGTCCGTCATTCGATGAGAGAGATTC AAAATGTGGGT 5269
    |||
OY 1161 CATATTGCAAAACCTTGGCCCTTGGGGCTCCACAGTCATGATGAGGCTCTCTCTGGCT 1220
    |||
```

```
DB 5270 CATATTGCAAAACCTTGGCCCTTGGGGCTCCACAGTCATGATGGGCTCTCTGGCT 5329
OY 1221 GCCACCACTGAGGCCCTTGTGTAATCTTTTCCGATGGATCCGGCTAAAGAAATAT 1280
    |||
DB 5330 GCCACCACTGAGGCCCTTGTGTAATCTTTTCCGATGGATCCGGCTAAAGAAATAT 5389
    |||
OY 1281 CGCGGATGAGGCTTCTCTCGATGCGATGACAGACCTCTGAGCCAGCAAGATATTC 1340
    |||
DB 5390 CGCGGATGAGGCTTCTCTCGATGCGATGACAGACCTCTGAGCCAGCAAGATATTC 5449
    |||

OY 1341 AG 1342
    ||
DB 5450 AG 5451

RESULT 16
LOCUS HUMIMPDH 6193 bp DNA linear PRI 22-MAY-1995
DEFINITION Homo sapiens (clone FFE-7) type II inosine monophosphate
dehydrogenase (IMPDH2) gene, exons 1-13, complete cds.
ACCESSION L33842
VERSION L33842.1 GI:602457
KEYWORDS
NAD-dependent; differentiation; inosine monophosphate
dehydrogenase; inosine-5'-monophosphate dehydrogenase; nucleotide
biosynthesis; proliferation associated gene.
Homo sapiens (tissue library: lambda GEM-11 (Promega)) blood DNA.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 6193)
REFERENCE
1 Glesne,D.A. and Huberman,E.
Cloning and sequence of the human type II IMP dehydrogenase gene
JOURNAL Blochem. Biophys. Res. Commun. 205 (1), 537-544 (1994)
MEDLINE 95091778
PUBMED 7999076
COMMENT Related sequences J04208 and I08114.
FEATURES
source 1. .6193
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="karyotype band P21.2-P24.2"
/tissue_type="blood"
/issue_11b="lambda GEM-11 (Promega)"
879. .1000
/promoter
/gene="IMPDH2"
/note="minimal promoter"
join(983..1182,1628..1676,1900..2001,2110..2184,
2516..2722,3385..3472,3546..3746,3825..3914,4014..4109,
5163..5451,5545..5688,5769..5852,5943..>6005)
/gene="IMPDH2"
join(983..1182,1628..1676,1900..2001,2110..2184,
2516..2722,3385..3472,3546..3746,3825..3914,4014..4109,
5163..5451,5545..5688,5769..5852,5943..>6005)
/gene="IMPDH2"
983..1182
/exon
/gene="IMPDH2"
/number=1
misc_feature
/gene="IMPDH2"
/note="major cap site"
1000
misc_feature
/gene="IMPDH2"
/note="minor cap site"
/gene="IMPDH2"
/note="minor cap site"
join(1085..1182,1628..1676,1900..2001,2110..2184,
2516..2722,3385..3472,3546..3746,3825..3914,4014..4109,
5163..5451,5545..5688,5769..5852,5943..5964)
/gene="IMPDH2"
/EC_number="1.1.1.205"
/codon_start=1
/product="inosine monophosphate dehydrogenase type II"
/protein_id="AAA67054.1"
/db_xref="GI:602458"
/translation="MADYLISGTSYVDDGLTAAQLFNCGDGLTYNDPLIPGVDF
```

intron	1183. .1627 /gene="IMPDH2" /number=1
exon	1628. .1676 /gene="IMPDH2" /number=2
intron	1677. .1899 /gene="IMPDH2" /number=2
exon	1900. .2001 /gene="IMPDH2" /number=3
intron	2002. .2109 /gene="IMPDH2" /number=3
exon	2110. .2184 /gene="IMPDH2" /number=4
intron	2185. .2515 /gene="IMPDH2" /number=4
exon	2516. .2722 /gene="IMPDH2" /number=5
intron	2723. .3384 /gene="IMPDH2" /number=5
exon	3385. .3472 /gene="IMPDH2" /number=6
intron	3473. .3545 /gene="IMPDH2" /number=6
exon	3546. .3746 /gene="IMPDH2" /number=7
intron	3747. .3824 /gene="IMPDH2" /number=7
exon	3825. .3914 /gene="IMPDH2" /number=8
intron	3915. .4013 /gene="IMPDH2" /number=8
exon	4014. .4109 /gene="IMPDH2" /number=9
intron	4110. .5162 /gene="IMPDH2" /number=9
exon	5163. .5451 /gene="IMPDH2" /number=10
intron	5452. .5544 /gene="IMPDH2" /number=10
exon	5545. .5688 /gene="IMPDH2" /number=11
intron	5689. .5768 /gene="IMPDH2" /number=11
exon	5769. .5852

Query Match	Best Local Similarity	14.6%	Score 242	DB 9	Length 6193
Matches 242	Conservative	0	Mismatches	0	Indels
		0	Gaps	0	
polyA-signal	1404 a	1597 c	1628 g	1564 t	
BASE COUNT	1404 a	1597 c	1628 g	1564 t	
ORIGIN					
intron	/gene="IMPDH2"	/number=12	5833	.5942	
exon	/gene="IMPDH2"	/number=12	5943	.>6005	
	/gene="IMPDH2"	/number=13	6000	.6005	
	/gene="IMPDH2"				
polyA-signal	/gene="IMPDH2"				
Query Match	14.6% <td>Score 242 <td>DB 9 <td>Length 6193 <td></td> </td></td></td>	Score 242 <td>DB 9 <td>Length 6193 <td></td> </td></td>	DB 9 <td>Length 6193 <td></td> </td>	Length 6193 <td></td>	
Best Local Similarity	100.0% <td>Pred. No. 2,2e-131 <td></td> <td></td> <td></td> </td>	Pred. No. 2,2e-131 <td></td> <td></td> <td></td>			
Matches 242 <td>Conservative</td> <td>0 <td>Mismatches</td> <td>0 <td>Indels</td> </td></td>	Conservative	0 <td>Mismatches</td> <td>0 <td>Indels</td> </td>	Mismatches	0 <td>Indels</td>	Indels
		0 <td>Gaps</td> <td>0 <td></td> </td>	Gaps	0 <td></td>	
Db 1101	GAGTATGCACGGCGCTTGTGTCCGCTCATCTGATGTGAGGAAATCAATGTGGGT	1160			
Db 5210	GAGTATGCACGGCGCTTGTGTGTCCGCTCATCTGATGTGAGGAAATCAATGTGGGT	5269			
Qy 1161	CATATTCGAAAGCCTTGTGGCCCTTGGGCGCTCCACAGTCATGATGAGGCTCTCTCGGCT	1220			
Db 5270	CATATTCGAAAGCCTTGTGGCCCTTGGGCGCTCCACAGTCATGATGAGGCTCTCTCGGCT	5329			
Qy 1221	GCCACACTGAGGCCCTGTGTGATATCTCTTTTCCGATGSGATCCGGCTAAAGAAATAT	1280			
Db 5330	GCCACACTGAGGCCCTGTGTGATATCTCTTTTCCGATGSGATCCGGCTAAAGAAATAT	5389			
Qy 1281	CGCGGTATGGGTCTCTCGATGCCATGTGACAAAGCACCACAGCCGCAAGACGATATTTC	1340			
Db 5390	CGCGGTATGGGTCTCTCGATGCCATGTGACAAAGCACCACAGCCGCAAGACGATATTTC	5449			
Qy 1341	AG 1342				
Db 5450	AG 5451				
RESULT 17					
HUMIMP					
LOCUS					
DEFINITION	Homo sapiens inosine monophosphate dehydrogenase type II gene,				
ACCESSION	L39210				
VERSION	L39210.1				
KEYWORDS	inosine monophosphate dehydrogenase; inosine monophosphate dehydrogenase type II.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 10557)				
TITLE	Zimmermann, A.G., Spychala, J. and Mitchell, B.S.				
JOURNAL	Characterization of the human inosine 5'-monophosphate dehydrogenase type II gene				
MEDLINE	J. Biol. Chem. 270 (12), 6808-6814 (1995)				
PUBMED	95204479				
REFERENCE	2 (bases 1 to 10557)				
AUTHORS	Zimmermann, A.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-NOV-1996) Department of Pharmacology, University of North Carolina, School of Medicine, CB7365, FLOB, Chapel Hill, NC 27599, USA				
REFERENCE	3 (bases 1 to 10557)				
AUTHORS	Zimmermann, A.G., Wright, K.L., Ting, J.P. and Mitchell, B.S.				
TITLE	Regulation of inosine 5'-monophosphate dehydrogenase type II gene expression in human T cells. Role for a novel 5' palindromic octamer sequence				
JOURNAL	J. Biol. Chem. 272 (36), 22913-22923 (1997)				
MEDLINE	97426442				
PUBMED	9278455				
COMMENT	On Dec 3, 1996 this sequence version replaced gi:777736.				



FEATURES  
source  
1.10557  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21.2-24.2"  
/cell\_type="Leukocyte"  
1509.1969  
join(1972.2069,2515.2563,2789.2890,2998.3072,3400.3606,4284.4351,4425.4624,4702.4792,4892.4987,6053.6196,6280.6424,6519.6662,6743.6826,6916.6990)  
/product="Inosine monophosphate dehydrogenase type II"  
join(1972.2069,2515.2563,2789.2890,2998.3072,3400.3606,4284.4351,4425.4624,4702.4792,4892.4987,6053.6196,6280.6424,6519.6662,6743.6826,6916.6937)  
/note="IMPDH"  
/codon\_start=1  
/product="Inosine monophosphate dehydrogenase type II"  
/protein\_id="AA070899.1"  
/db\_xref="GI:1702964"  
/translation="MADYLIISGTSYVDDGLTAAQIFNCGDGLTYNDFLLPGYIDF  
TADQVLTSLTKITLTKPLVSSPMDFVTEAGMAIALMTGIGFIHNCIPPEOAN  
EVRKYKEOGFITDPVLSKPDVRDVEAKARQFICPITDGRMGSRLLVGISS  
RIDFLKEEHDCFLIELMTKREDLVVAPAGITLKEAKEYILORSKKGKPIYNEDDL  
VAIARLTKKRDYDPLASKKOLLGCAIGTHEDKRYRDLIAQACVDVYVLDSS  
OCNSTFOIMIKYITDKYIPNDVIGGVVYTAQAKNLDAGVDALRAGMSGICITQ  
EVLACGPQATAVVSEYARREGVPIADGIIQNVGLAKALAGASTVMGSLIAA  
TTEAPGEFFSDGIRLKYRCMGSLDAMDHLSSONKFFSEADKIKVAGGVGAODK  
GSIHKFVPLYLAGIIOHSCODIGAKSLTOVRAMVSGEAKFEKRTSSAOGEGVSHLS  
YEKRLF"

BASE COUNT 2348 a 2852 c 2963 g 2394 t  
ORIGIN

Query Match 12.6%; Score 209; DB 9; Length 10557;  
Best Local Similarity 100.0%; Pred. No. 8.7e-112;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 AGAAATATACAGGATTCATCACAGACCCCTGTCTCTCACCCCA GGATCGGCTGC 429  
DB 3398 AGAAATATACAGGATTCATCACAGACCCCTGTCTCTCACCCCA GGATCGGCTGC 3457  
QY 430 GGGATGTTTTAGAGCCAGCCCGCATGTTTCGCGTATCCCA CACAGACACAG 489  
DB 3458 GGGATGTTTTAGAGCCAGCCCGCATGTTTCGCGTATCCCA CACAGACACAG 3517  
QY 490 GCGGATGGGAGCCGCTGTCGGCATCATCTCCACAGGACATTG TTTTCTCAAG 549  
DB 3518 GCGGATGGGAGCCGCTGTCGGCATCATCTCCACAGGACATTG TTTTCTCAAG 3577  
QY 550 AGGAGACATGACTGTTCTTGGAGAG 578  
DB 3578 AGGAGACATGACTGTTCTTGGAGAG 3606

RESULT 18  
BC009321 1851 bp mRNA linear PRI 12-JUL-2001  
LOCUS Homo sapiens, clone MGC:16650 IMAGE:4123521, mRNA, complete cds.  
ACCESSION BC009321  
VERSION BC009321.1 GI:14424584  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1851)  
Strausberg,R.  
Direct Submission  
Submitted (12-JUN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [ccgabs-remail.nih.gov](mailto:ccgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Shewchenko,Y., Weltherby,K.D., Beckstrom-Sternberg,S.M.,  
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,  
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,  
Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C.,  
McDowell,J., Pearson,R., Snyder,B., Stancirip,S., Thomas,P.J.,  
Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,  
Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 26 Row: e Column: 10.  
Location/Qualifiers  
1.1851  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:16650 IMAGE:4123521"  
/tissue\_type="Muscle, rhabdomyosarcoma"  
/clone\_lib="NIH-MGC-17"  
/lab\_host="DH10B-R"  
/note="vector: pOTB7"  
191..688  
/codon\_start=1  
/product="Unknown (protein for MGC:16650)"  
/protein\_id="AAH09321.1"  
/db\_xref="GI:14424585"  
/translation="MCYTHALFLQRLPRRSEAGRPRLRPVSVARSVATKSAG  
LYMRPDRGLGAVLGPTRRCPIYLAAMDYLSGTSYVPPDGLTAAQDLFNCGLDLY  
KCGPGLTMAKGAAPWCKOMGASESRILGRGGVSGVIGHPPLTLRIHLRCLGY  
SEASQ"

BASE COUNT 465 a 551 c 400 g 435 t  
ORIGIN

Query Match 8.3%; Score 137; DB 9; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 6.4e-69;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GCGGCTCTGGAGACAGCGCGGTCTGTGTTGGCATGGCGACTACCTGATTAGT 68  
DB 362 GCGGCTCTGGAGACAGCGCGGTCTGTGTTGGCATGGCGACTACCTGATTAGT 421  
QY 69 GGGGGACAGCTCTAGCTGCAGACGAGGACTCAGACAGCAGCTCTTCAACGCGGA 128  
DB 422 GGGGGACAGCTCTAGCTGCAGACGAGGACTCAGACAGCAGCTCTTCAACGCGGA 481  
QY 129 GACGCGCTCAGCTTACAA 145  
DB 482 GACGCGCTCAGCTTACAA 498

RESULT 19  
HU01MPDEH 360 bp DNA linear PRI 12-JUN-1993  
LOCUS Homo sapiens, clone MGC:16650 IMAGE:4123521, mRNA, complete cds.  
DEFINITION Human inosine 5'-monophosphate dehydrogenase subunit II gene, last  
3 exons.  
L08114  
108114.1 GI:292239  
ACCESSION L08114  
VERSION L08114.1 GI:292239  
KEYWORDS imp dehydrogenase; inosine 5'-monophosphate dehydrogenase  
dehydrogenase.  
SOURCE Homo sapiens DNA.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (sites)  
 AUTHORS Glesne, D., Collart, F., Varkony, T., Drabkin, H., and Huberman, E.  
 TITLE Chromosomal localization and structure of the human type II IMP  
 dehydrogenase gene (IMPDH2)  
 JOURNAL Genomics 16 (1), 274-277 (1993)  
 MEDLINE 93253398  
 PUBMED 8098009

FEATURES  
 source Location/Qualifiers  
 1. 360  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="3p21.2-p24.2"  
 /cell\_line="HL-605"  
 join<1. .59,140. .223,314. .335)  
 /EC\_number="1.1.1.205"  
 /standard\_name="IMPDH2"  
 /note="subunit II"  
 /codon\_start=1  
 /product="inosine monophosphate dehydrogenase"  
 /protein\_id="AAA36113.1"  
 /db\_xref="GI:292240"  
 /translation="TAGIOHSCODIGAKSLTQVRAMVSGELKFEKRTSSAOVEGVH  
 SHSYEKRLF"

CDS  
 intron  
 exon  
 intron  
 exon  
 3'UTR  
 BASE COUNT 70 a 106 c 99 g 85 t  
 ORIGIN

Query Match  
 Best Local Similarity 100.0%; Pred. No. 9.9e+40;  
 Matches 88: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1486 GAGCCATGATGTACTCTGGGAGCTTAACTTTGAGAAGAACGCTCAGCCAGGTGG 1545  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 139 GAGCATGATGTACTCTGGGAGCTTAACTTTGAGAAGAACGCTCAGCCAGGTGG 198

QY 1546 AAGTGGCGTCATACGCTTCATTCGTA 1573  
 ||||||||||||||||||||||||||||||||||||  
 Db 199 AAGTGGCGTCATACGCTTCATTCGTA 226

RESULT 20  
 CROUTMP 1614 bp mRNA linear ROD 10-JUN-1993  
 LOCUS Chinese hamster inosine-5'-monophosphate dehydrogenase (IMP) mRNA,  
 DEFINITION complete cds.  
 ACCESSION J04209  
 VERSION J04209.1 GI:191119  
 KEYWORDS inosine-5'-monophosphate dehydrogenase.  
 SOURCE Chinese hamster cell line E29Pro+, clone CIMP.  
 ORGANISM Crictetus griseus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Crictetinae;  
 Crictetus  
 1 (bases 1 to 1614)  
 Collart, F.R. and Huberman, E.  
 TITLE Cloning and sequence analysis of the human and Chinese hamster  
 inosine-5'-monophosphate dehydrogenase cDNAs  
 JOURNAL J. Biol. Chem. 263 (30), 15769-15772 (1988)  
 MEDLINE 89008491  
 PUBMED 2902093

COMMENT  
 Draft entry and computer-readable sequence for [1] kindly provided  
 by F. Collart, 07-SEPT-1988.  
 FEATURES  
 source Location/Qualifiers  
 1. 1614  
 /organism="Crictetus griseus"  
 /db\_xref="taxon:10029"  
 1. 1614  
 gene

mrna  
 CDS  
 /gene="IMP"  
 <1. .1614  
 /gene="IMP"  
 26. .1570  
 /gene="IMP"  
 /EC\_number="1.1.1.205"  
 /codon\_start=1  
 /product="inosine-5'-monophosphate dehydrogenase"  
 /protein\_id="AAA36993.1"  
 /db\_xref="GI:304517"  
 /translation="MADYLISGCTSYVDDCLTAQQLFNCGDGLTYNDPLIPGYIDF  
 TADQVLTSLTKTKITLTKPLVSSPMDVTETEGMAIAMLTGIGLGFHHNCTPERQAN  
 EVRKVKRYEGFTDPVVLSPDRVDVFEAKARFCGIPITDTGRMSRLVGIIS  
 RDIDFLKEEHRFLFLEIMTKREDLVAPAGITLKEANIDILORSKGLPIYNENDEL  
 VAIARTLKKRNDVPLASDKAKKQLLGAAIGTHEDDKYRLDLLALAGVYVYIDSS  
 OGNSTFQINMKIKYMEKYPNIOYIGNVYTAQAKNLIDACVDALRVGMSGCICITQ  
 EYLACRPQATRAVIVSEIARFEGVYADGILQVNGHIAKALAGASTVMGSLAA  
 TTEARGEFESDGIKLRKRYRGMGLDAMDKHLSSONRYFSEADKIKVAQVGSAYQDK  
 GSIHKFVFPYLLAGIOHSCODIGAKSLTQVRAMVSGELKFEKRTSSAOVEGVSHLS  
 YEKRLF"

BASE COUNT 417 a 369 c 421 g 407 t  
 ORIGIN 2 bp upstream of Mui site.

Query Match  
 Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
 Matches 62: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GGTATTGGCTTATCCACCAACGTATACCTGATTCAGGCCAATGAAGTTGGAAA 365  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 284 GGTATTGGCTTATCCACCAACGTATACCTGATTCAGGCCAATGAAGTTGGAAA 343

QY 366 GT 367  
 ||  
 Db 344 GT 345

RESULT 21  
 ARI41890 1620 bp DNA linear PAT 08-AUG-2001  
 LOCUS ARI41890  
 DEFINITION Sequence 2 from patent US 6147194.  
 ACCESSION ARI41890  
 VERSION ARI41890.1 GI:15101406  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1 (bases 1 to 1620)  
 Collart, F.R. and Huberman, E.  
 TITLE Eukaryotic IMPDH polynucleotide and antibody compositions  
 JOURNAL Patent: US 6147194-A 2 14-NOV-2000;  
 FEATURES  
 source Location/Qualifiers  
 1. 1620  
 /organism="unknown"  
 BASE COUNT 423 a 369 c 421 g 407 t  
 ORIGIN

Query Match  
 Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
 Matches 62: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GGTATTGGCTTATCCACCAACGTATACCTGATTCAGGCCAATGAAGTTGGAAA 365  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 284 GGTATTGGCTTATCCACCAACGTATACCTGATTCAGGCCAATGAAGTTGGAAA 343

QY 366 GT 367  
 ||  
 Db 344 GT 345

RESULT 22  
 106572 1620 bp DNA linear PAT 02-DEC-1994  
 LOCUS 106572  
 DEFINITION Sequence 3 from Patent WO 9001545.

```

ACCESSION 106572
VERSION 106572.1 GI:589654
KEYWORDS
SOURCE unknown
ORGANISM unknown
REFERENCE 1 (bases 1 to 1620)
AUTHORS Collart,F.R. and Huberman,E.
TITLE METHOD AND MATERIALS RELATING TO IMPDH AND GMP
JOURNAL Patent: WO 9001545-A 3 22-FEB-1990;
FEATURES
    source
        1..1620
            /organism="unknown"
BASE COUNT 423 a 369 c 421 g 407 t
ORIGIN
Query Match 3.7%: Score 62; DB 6; Length 1620;
Best Local Similarity 100.0%; Pred No.2.6e-24;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 GGATATGGCTTCATCCACCACTGTACACCTGGAATTCAGGCCAATAAGTTGGGAAA 365
    |||||||
DB 284 GGATATGGCTTCATCCACCACTGTACACCTGGAATTCAGGCCAATAAGTTGGGAAA 343
QY 366 GT 367
DB 344 GT 345

RESULT 23
MUSIMPDA 1545 bp mRNA linear ROD 27-JUL-1994
LOCUS Mus musculus IMP dehydrogenase (IMPD) mRNA, complete cds.
DEFINITION M98333.1 GI:425157
ACCESSION M98333.1 GI:425157
VERSION IMP dehydrogenase.
KEYWORDS Mus musculus brain cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1545)
AUTHORS Lightfoot,T. and Snyder,F.F.
TITLE Gene amplification and dual point mutations of mouse IMP
dehydrogenase associated with cellular resistance to mycophenolic
acid
JOURNAL Biochim. Biophys. Acta 1217 (2), 156-162 (1994)
MEDLINE 94153991
PUBMED 7905545
FEATURES
    source
        Location/Qualifiers
            1..1545
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /cell_type="neuroblastoma"
                /tissue_type="brain"
            1..1545
                /gene="IMPD"
            1..1545
                /gene="IMPD"
                /product="IMP dehydrogenase"
                /codon_start=1
                /protein_id="AA020181.1"
                /db_xref="GI:425158"
                /translation="MADYLSISGTSVPPDGLTAQQLNCGDGLTYNDPLILPGYIDF
                TADVDLTALTKKTLTKPLVSSPMDYTEGMATATAATGGIGFIHNCTPEFOAN
                EVRKVKTEOGFITDPVLSPKRDVADVEAKARHGFCTPTDTGRKMSRLVGIISS
                RIDPLKEEHRLEEMTKREDLVNAPAGVTLKEAEILQSKSKGKLPYVNEDEL
                VAIARTDLKKNRDPPLASKDKQLLCAAGTTHEDKYRIDLLALACVDVVDLDS
                ONSIFQIMIKYIKKKYPSLOVIGNVVTAQAOKNL DAGVDALRVSNGSSICITQ
                EVLACGPQATAVYKVEYARRGVPIADGCIQNGE IAKALALGASTVVMGSLIAA
                TTEARPEYFSDGIRLKRYKRGMSLDAMDKHLSSONR ESEADKIKVAGCGSAVQDK
                GSIIHKVPYFLIANGIOHSCDIDAKSLITVRAAMVSGELKFEKRTSSAOVGCVHLSHS
                YKRLF"
            998
variation
998

```

```

variation
1052
/organism="Mus musculus"
/feature="IMPD"
/feature="a in mutation"
BASE COUNT 390 a 374 c 419 g 362 t
ORIGIN
Query Match 3.4%: Score 56; DB 10; Length 1545;
Best Local Similarity 100.0%; Pred No.9.5e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 777 CTGCTGTGTGGCAGCATTGGCAGCTCATGAGATGACAACTATAGCTGACTT 832
    |||||||
DB 730 CTGCTGTGTGGCAGCATTGGCAGCTCATGAGATGACAACTATAGCTGACTT 785

RESULT 24
BC010314 1680 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, inosine 5'-phosphate dehydrogenase 2, clone MGC:6193
ACCESSION IMAGE:2646679, mRNA, complete cds.
VERSION BC010314
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1680)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: mgc-help@nci.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Toon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Plate: 5 Row: C Column: 15.
FEATURES
    source
        Location/Qualifiers
            1..1680
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /map="CZCH II"
                /clone="MGC:6193 IMAGE:2646679"
                /tissue_type="Mammary tumor metastasized to lung. Tumor
                arose spontaneously from a senescent normal mammary
                (clonal) outgrowth infected with the virus MMV."
                /lab_host="NC1 CGAP Lu29"
                /note="Vector: pcMV-SPOrt6"
            34..1578
                /codon_start=1
                /product="inosine 5'-phosphate dehydrogenase 2"
                /protein_id="AA10314.1"
                /db_xref="GI:16307531"
                /db_xref="LocusID:23918"
                /translation="MADYLSISGTSVPPDGLTAQQLFNCGDGLTYNDPLILPGYIDF

```

TADVDLTSLTKITLKTPLVSSPMDVTEAGMATAMLTGIGEIFHNCTPEOAN  
EVRKVKYEOGFIITDEPVLSPPDRVDEVEAKRHGFCGIPITDTRMGSRIVGIIS  
RPIDLKEHEHDFLEIEMTKREDLVANAGVITKIANLILORSKRGKPIYINDEL  
VALIARTDKKRNNDVPLASKDKKQILCGAIGTHEDDDYRIDLLALAGVIVVDDSS  
QGNSTFQIMWITIKETIPSLQVIGGNVTAQAOKNLIDAGVDALFVGMGSGSICITQ  
EVLACGRPOATAVYKSEYARFVGVIDGGIONGHITAKALALGASTVMGSLTAA  
TEVAREYFPEFSDGIRLKRKGMSLDAMOKHLSQNRYPSEADKIVAGVSGAODK  
GSIHKFVPLVINGIOHSCODIGAKSLTQVRAMVYSGELKFEKRTSSAQVEGVHLSHS  
YKRLFF"

BASE COUNT 465 a 396 c 435 g 384 t

ORIGIN

Query Match 3.4%; Score 56; DB 10; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 9.5e-21;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 777 CTGCTGTGTGGGAGCCATTGCACATGAGATGACAGATATAGCTGACTT 832  
DB 763 CTGCTGTGTGGGAGCCATTGCACATGAGATGACAGATATAGCTGACTT 818

RESULT 25  
MUSIMPD 1707 bp mRNA linear ROD 12-JUN-1993  
LOCUS Mouse IMP dehydrogenase mRNA, complete cds.  
DEFINITION M33934  
ACCESSION M33934.1 GI:198393  
VERSION IMP dehydrogenase.  
KEYWORDS Mouse, cDNA to mRNA.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1707)  
AUTHORS Tiedeman, A.A. and Smith, J.M.  
TITLE Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase  
JOURNAL Gene 97 (2), 289-293 (1991)  
MEDLINE 91153661  
PUBMED 1671845  
COMMENT Draft entry and computer-readable sequence for [unpublished (1990)]  
kindly submitted  
by J.M.Smith 27-APR-1990.  
Author address:  
Dr. J.M. Smith  
Seattle Biomedical Research Institute  
4 Nickerson St.  
Seattle, WA 98109.  
FEATURES  
source  
Location/Qualifiers  
1..1707  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
67..1611  
/note="IMP dehydrogenase (EC 1.2.1.14)"  
/codon\_start=1  
/protein\_id="AAA9311.1"  
/db\_xref="GI:309413"  
/translation="MADYLIISGTSYVPDDGLTAQOLFNCGGDLTYNDLILPGYIDF  
TADVDLTSLTKITLKTPLVSSPMDVTEAGMATAMLTGIGEIFHNCTPEOAN  
EVRKVKYEOGFIITDEPVLSPPDRVDEVEAKRHGFCGIPITDTRMGSRIVGIIS  
RPIDLKEHEHDFLEIEMTKREDLVANAGVITKIANLILORSKRGKPIYINDEL  
VALIARTDKKRNNDVPLASKDKKQILCGAIGTHEDDDYRIDLLALAGVIVVDDSS  
QGNSTFQIMWITIKETIPSLQVIGGNVTAQAOKNLIDAGVDALFVGMGSGSICITQ  
EVLACGRPOATAVYKSEYARFVGVIDGGIONGHITAKALALGASTVMGSLTAA  
TEVAREYFPEFSDGIRLKRKGMSLDAMOKHLSQNRYPSEADKIVAGVSGAODK  
GSIHKFVPLVINGIOHSCODIGAKSLTQVRAMVYSGELKFEKRTSSAQVEGVHLSHS  
YKRLFF"

BASE COUNT 457 a 408 c 451 g 391 t

ORIGIN

Query Match 3.4%; Score 56; DB 10; Length 1707;  
Best Local Similarity 100.0%; Pred. No. 9.4e-21;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 777 CTGCTGTGTGGGAGCCATTGCACATGAGATGACAGATATAGCTGACTT 832

DB 796 CTGCTGTGTGGGAGCCATTGCACATGAGATGACAGATATAGCTGACTT 851

RESULT 26  
AC084743/c  
LOCUS  
DEFINITION  
AC084743 182658 bp DNA linear HTG 03-FEB-2001  
Mus musculus clone RP23-104L6, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
unordered pieces.  
AC084743  
AC084743.1 GI:11138180  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Mus musculus.  
MUSIMPD  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 182658)  
AUTHORS McCombie, W.R., Baker, J.P., Bahrel, A., Bal, H., Ballja, V.,  
Dedhia, N.N., de la Bastide, M., Huang, E.N., King, L., Kirchoff, K.A.,  
Miller, B., Nascimento, L.U., O'Shaughnessy, A.L., Preston, R.R.,  
Rodriguez, S., Santos, L., Shah, R.S., Spiegel, L.A., Toth, K., Vill, M.D.  
and Zutavern, T.  
TITLE Mouse Genomic Sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 182658)  
AUTHORS McCombie, W.R.  
TITLE Direct Submission  
JOURNAL Submitted (11-NOV-2000) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
COMMENT  
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor  
Laboratory  
Center code: CSHL  
Web site: <http://www.cshl.org/genseq>  
Contact: [mccombie@cshl.org](mailto:mccombie@cshl.org)  
Project Information  
Center project name: RP23-104L6  
Center clone name: RP23-104L6  
\*\*\*\*\*  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 60 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
9239 9238: contig of 9238 bp in length  
9350 9349: gap of unknown length  
9350 17009: contig of 7660 bp in length  
17010 17120: gap of unknown length  
17121 23426: contig of 6306 bp in length  
23427 23537: gap of unknown length  
23538 29489: contig of 5952 bp in length  
29490 29600: gap of unknown length  
29601 35374: contig of 5774 bp in length  
35375 35485: gap of unknown length  
35486 40978: gap of 5433 bp in length  
40979 41089: gap of unknown length  
41090 46554: contig of 5465 bp in length  
46555 46665: gap of unknown length  
46666 51860: contig of 5195 bp in length  
51861 51971: gap of unknown length  
51972 57074: contig of 5103 bp in length  
57075 57185: gap of unknown length  
57186 61328: contig of 4143 bp in length  
61329 65216: gap of unknown length  
65216 65327: contig of 3777 bp in length  
65327 65328: gap of unknown length  
65328 68892: contig of 3365 bp in length  
68893 72554: gap of unknown length  
72554 72554: contig of 3551 bp in length

```
* 72555 72665: gap of unknown length
* 72666 76002: contig of 3337 bp in length
* 76003 76113: gap of unknown length
* 76114 79430: contig of 3317 bp in length
* 79431 79541: gap of unknown length
* 79542 82808: contig of 3267 bp in length
* 82809 82919: gap of unknown length
* 82920 86165: contig of 3246 bp in length
* 86166 86276: gap of unknown length
* 86277 89403: contig of 3127 bp in length
* 89404 89514: gap of unknown length
* 89515 92509: contig of 3095 bp in length
* 92610 92720: gap of unknown length
* 92721 95799: contig of 3079 bp in length
* 95800 95910: gap of unknown length
* 95911 98749: contig of 2839 bp in length
* 98750 98860: gap of unknown length
* 98861 101564: contig of 2704 bp in length
* 101565 101675: gap of unknown length
* 101676 104281: contig of 2606 bp in length
* 104282 104392: gap of unknown length
* 104393 106983: contig of 2591 bp in length
* 106984 107094: gap of unknown length
* 107095 109664: contig of 2570 bp in length
* 109665 109775: gap of unknown length
* 109776 112176: contig of 2401 bp in length
* 112177 112287: gap of unknown length
* 112288 114642: contig of 2355 bp in length
* 114643 114753: gap of unknown length
* 114754 117092: contig of 2339 bp in length
* 117093 117203: gap of unknown length
* 117204 119538: contig of 2335 bp in length
* 119539 119649: gap of unknown length
* 119650 121973: contig of 2324 bp in length
* 121974 122064: gap of unknown length
* 122065 124406: contig of 2322 bp in length
* 12407 124517: gap of unknown length
* 124518 126788: contig of 2271 bp in length
* 126789 126899: gap of unknown length
* 126900 129169: contig of 2270 bp in length
* 129170 129280: gap of unknown length
* 129281 131544: contig of 2264 bp in length
* 131545 131655: gap of unknown length
* 131656 133860: contig of 2205 bp in length
* 133861 133971: gap of unknown length
* 133972 136158: contig of 2187 bp in length
* 136159 136269: gap of unknown length
* 136270 138445: contig of 2176 bp in length
* 138446 138556: gap of unknown length
* 138557 140663: contig of 2107 bp in length
* 140664 140774: gap of unknown length
* 140775 142854: contig of 2080 bp in length
* 142855 142965: gap of unknown length
* 142966 145037: contig of 2072 bp in length
* 145038 145148: gap of unknown length
* 145149 147178: contig of 2030 bp in length
* 147179 147289: gap of unknown length
* 147290 149310: contig of 2021 bp in length
* 149311 149421: gap of unknown length
* 149422 151406: contig of 1985 bp in length
* 151407 151517: gap of unknown length
* 151518 153491: contig of 1974 bp in length
* 153492 153602: gap of unknown length
* 153603 153645: contig of 1943 bp in length
* 153646 153656: gap of unknown length
* 153657 157584: contig of 1928 bp in length
* 157585 157694: gap of unknown length
* 157695 159571: contig of 1877 bp in length
* 159572 159681: gap of unknown length
* 159682 161531: contig of 1850 bp in length
* 161532 161641: gap of unknown length
* 161642 163442: contig of 1801 bp in length
* 163443 163552: gap of unknown length

* 163553 165331: contig of 1779 bp in length
* 165332 165441: gap of unknown length
* 165442 167193: contig of 1752 bp in length
* 167194 167303: gap of unknown length
* 167304 169032: contig of 1729 bp in length
* 169033 169142: gap of unknown length
* 169143 170851: contig of 1709 bp in length
* 170852 170961: gap of unknown length
* 170962 172660: contig of 1699 bp in length
* 172661 172770: gap of unknown length
* 172771 174467: contig of 1697 bp in length
* 174468 174577: gap of unknown length
* 174578 176236: contig of 1659 bp in length
* 176237 176346: gap of unknown length
* 176347 177973: contig of 1627 bp in length
* 177974 178083: gap of unknown length
* 178084 179692: contig of 1609 bp in length
* 179693 179802: gap of unknown length
* 179803 181384: contig of 1582 bp in length
* 181385 181494: gap of unknown length
* 181495 182658: contig of 1164 bp in length.

FEATURES
    source          1..182658
                    /organism="Mus musculus"
                    /db_xref="taxon:10090"
                    /clone="RP23-104L6"

BASE COUNT      45897 a 42185 c 39609 g 48199 t 6768 others
ORIGIN
Query Match          3.4%: Score 56; DB 2: Length 182658;
Best Local Similarity 100.0%; Pred. No. 6.8e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 777 CTGCTGTGTGGGCGACCATTTGGCACTCATGAGATGACAGTATAGCTGCACTT 832
    |||||||
Db 147783 CTGCTGTGTGGGCGACCATTTGGCACTCATGAGATGACAGTATAGCTGCACTT 147728

RESULT 27
AC122514
LOCUS          225727 bp DNA linear HTG 30-JUL-2002
DEFINITION    Mus musculus chromosome UNK clone RP24-455L10, WORKING DRAFT
ACCESSION     AC122514
VERSION       AC122514.2 GI:22004610
KEYWORDS      HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 225727)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE         The sequence of Mus musculus clone
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 225727)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE         Direct Submission
JOURNAL       Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE     3 (bases 1 to 225727)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE         Direct Submission
JOURNAL       Submitted (30-JUL-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
COMMENT       On Jul 30, 2002 this sequence version replaced gi:21105978.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BR0455L10
```

```
----- Summary Statistics -----
Sequencing vector: M13: 0%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembley program: Phrap; version 0.990319
Consensus quality: 224286 bases at least Q40
Consensus quality: 224589 bases at least Q30
Consensus quality: 224737 bases at least Q20
Insert size: 174000; agarose-1p
Insert size: 225627; sum-of-ctrls
Quality coverage: 15.05 in Q20 bases; agarose-1p
Quality coverage: 11.65 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 53314: contig of 53314 bp in length
* 53315 53414: gap of unknown length
* 53415 225727: contig of 172313 bp in length.
FEATURES
  source
    1. 225727
       Location/Qualifiers
         /organism="Mus musculus"
         /db_xref="taxon:10090"
         /chromosome="X"
         /clone="RP24-455L10"
         1. 53314
            /note="assembly_name:Contig10"
            misc_feature
              53415..225727
                 /note="assembly_name:Contig11"
BASE COUNT  70387 a 42350 c 42349 g 70510 t 131 others
ORIGIN
Query Match      3.4%; Score 56; DB 2; Length 225727;
Best Local Similarity 100.0%; Pred. No. 6.7e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 777 CTGCTGTGTGGGACGACCTGACCTCATGAGATGACAGATGAGCTGACTT 832
      ||||||||||||||||||||||||||||||||||||||||||||||||
Db 110095 CTGCTGTGTGGGACGACCTGACCTCATGAGATGACAGATGAGCTGACTT 110150

RESULT 28
LOCUS      HS86C2R      332 bp      DNA      linear      PRI 23-OCT-1995
DEFINITION H.sapiens Cpg island DNA genomic Msel fragment, clone 86c2, reverse
            read cpg86c2.rtl.a.
ACCESSION  Z66268
VERSION    Z66268.1 GI:1039090
KEYWORDS  Cpg island; genomic Msel fragment.
SOURCE    Homo sapiens.
           Homo sapiens
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 332)
AUTHORS  Macdonald,M., Huckle,E., Wilkinson,P. and Mickleam,G.
JOURNAL   Direct Submission
          Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
          CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
          2 (bases 1 to 332)
          Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
          Purification of Cpg islands using a methylated DNA binding column
          Nat. Genet. 6 (3), 236-244 (1994)
REFERENCE  8012384
JOURNAL   MEDLINE
          94282070
          PUBMED
          8012384
COMMENT   Vector: pGEM-52f(-)
          Clones are available from the UK MRC Human Genome Mapping Project
```

```
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hmp.mrc.ac.uk/ for details
or contact: biohelp@hmp.mrc.ac.uk.
-----
FEATURES
  source
    1. 332
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="86c2"
       /sex="male"
       /tissue_type="blood"
       /clone_11b="CGI-1"
       /dev_stage="adult"
BASE COUNT  51 a 100 c 119 g 62 t
ORIGIN
Query Match      3.2%; Score 53; DB 9; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.4e-19;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 GACACCGCGCGGTCTGTTGGCCATGCGCGACTACTGATTAGTGGGG 73
      ||||||||||||||||||||||||||||||||||||||||||||||||
Db 260 GACACCGCGCGGTCTGTTGGCCATGCGCGACTACTGATTAGTGGGG 312

RESULT 29
LOCUS      AX190186/c      51 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 365 from Patent WO0147942.
ACCESSION  AX190186
VERSION    AX190186.1 GI:15143562
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 51)
AUTHORS  Shinkets,R.A. and Leach,M.
JOURNAL   Nucleic acids containing single nucleotide polymorphisms and
          methods of use thereof
          Patent: WO 0147942-A 365 05-JUL-2001;
          Curagen Corporation (US)
FEATURES
  source
    1. 51
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /note="1 of 2 allelic variants (366 is other
          entry)-Accession number cg43922096"
BASE COUNT  12 a 20 c 13 g 6 t
ORIGIN
Query Match      3.1%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 27 CGCGGCTGTCTGTGTTGGCCATGCGCGACTACTGATTAGTGGGGCAGC 77
      ||||||||||||||||||||||||||||||||||||||||||||||||
Db 51 CGCGGCTGTCTGTGTTGGCCATGCGCGACTACTGATTAGTGGGGCAGC 1

RESULT 30
LOCUS      ARI41893/c      393 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6147194.
ACCESSION  ARI41893
VERSION    ARI41893.1 GI:15101409
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
           Unclassified.
REFERENCE  1 (bases 1 to 393)
AUTHORS  Collart,F.R. and Huberman,E.
JOURNAL   Eukaryotic IMPDH polynucleotide and antibody compositions
          Patent: US 6147194-A 5 14-NOV-2000;
```

FEATURES  
source  
BASE COUNT 99 a 111 c 89 g 94 t  
ORIGIN

Query Match 2.5%; Score 41; DB 6; Length 393;  
Best Local Similarity 100.0%; Pred. No. 8.7e-12;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 843 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 883  
Db 322 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 282

RESULT 31  
LOCUS AR141894 393 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 6 from patent US 6147194.  
ACCESSION AR141894  
VERSION AR141894.1 GI:15101410  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1 (bases 1 to 393)  
AUTHORS Collart,F.R. and Huberman,E.  
TITLE Eukaryotic IMPDH polynucleotide and antibody compositions  
JOURNAL Patent: US 6147194-A 6 14-NOV-2000;  
FEATURES  
source 1..393  
BASE COUNT 94 a 89 c 111 g 95 t  
ORIGIN

Query Match 2.5%; Score 41; DB 6; Length 393;  
Best Local Similarity 100.0%; Pred. No. 8.7e-12;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 843 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 883  
Db 72 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 112

RESULT 32  
LOCUS 106574 393 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 5 from Patent WO 9001545.  
ACCESSION 106574  
VERSION 106574.1 GI:589656  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1 (bases 1 to 393)  
AUTHORS Collart,F.R. and Huberman,E.  
TITLE METHOD AND MATERIALS RELATING TO IMPDH AND GMP PRODUCTION  
JOURNAL Patent: WO 9001545-A 5 22-FEB-1990;  
FEATURES  
source 1..393  
BASE COUNT 99 a 111 c 89 g 94 t  
ORIGIN

Query Match 2.5%; Score 41; DB 6; Length 393;  
Best Local Similarity 100.0%; Pred. No. 8.7e-12;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 843 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 883  
Db 322 GCTGCTGATGATGATGCTTTGGACCTCTCCACAGGAAA 282

RESULT 33  
LOCUS AC107280/C 114558 bp DNA linear HTG 13-JUL-2002  
DEFINITION Rattus norvegicus clone CH230-81H4, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
AC107280  
AC107280.3 GI:21737470  
HTG: HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

REFERENCE  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alsprouks,S.L., Amaratunge,H.C., Are,J.R., Ayala,M., Banks,T., Barbaita,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D., Bouck,U., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C., Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douhwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulik,S., Hume,D., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H., Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,F., Maxhiney,F., Mcleod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S., Ogum,H., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshari,N., Stinson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,S., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (18-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 114558)  
Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18702062.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>

```
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPRM
Center clone name: CH230-81H4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 59289 bases at least Q40
Consensus quality: 62371 bases at least Q30
Consensus quality: 65704 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 58 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1407: contig of 1407 bp in length
1408
1507: gap of unknown length
1508
2785: contig of 1278 bp in length
2786
2885: gap of unknown length
2886
4306: contig of 1421 bp in length
4307
4406: gap of unknown length
4407
5791: contig of 1385 bp in length
5792
5891: gap of unknown length
5892
7219: contig of 1328 bp in length
7220
7319: gap of unknown length
7320
8441: contig of 1122 bp in length
8442
8541: gap of unknown length
8542
9663: contig of 1422 bp in length
9664
10063: gap of unknown length
10064
11368: contig of 1305 bp in length
11369
11468: gap of unknown length
11469
12924: contig of 1456 bp in length
12925
13024: gap of unknown length
13025
14074: contig of 1050 bp in length
14075
14174: gap of unknown length
14175
15767: contig of 1593 bp in length
15768
15867: gap of unknown length
15869
17131: contig of 1264 bp in length
17132
17231: gap of unknown length
17232
18519: contig of 1288 bp in length
18520
18619: gap of unknown length
18620
19671: contig of 1052 bp in length
19672
19771: gap of unknown length
21140: contig of 1369 bp in length
21141
21240: gap of unknown length
21241
22449: contig of 1209 bp in length
22450
22549: gap of unknown length
22550
23904: contig of 1355 bp in length
23905
24004: gap of unknown length
24005
25150: contig of 1146 bp in length
25151
25250: gap of unknown length
25251
26269: contig of 1019 bp in length
26270
26369: gap of unknown length
26370
27985: contig of 1616 bp in length
27986
28085: gap of unknown length
28407: gap of 1322 bp in length
29507: gap of unknown length
29508
30634: contig of 1127 bp in length
30635
30734: gap of unknown length
30735
31927: contig of 1193 bp in length
31928
32027: gap of unknown length
33028
33335: contig of 1308 bp in length
33336
33435: gap of unknown length
33436
35511: contig of 2076 bp in length
35512
37208: contig of 1597 bp in length
```

```
37209
37309
38952: contig of 1644 bp in length
38953
39053
40524: contig of 1472 bp in length
40525
40624: gap of unknown length
42195
42294: gap of 1570 bp in length
42295
42382: gap of unknown length
42383
43682: gap of 1488 bp in length
43684
43883
45446: gap of unknown length
45447
45546: gap of 1564 bp in length
45447
45546: gap of unknown length
45547
46984: contig of 1438 bp in length
46985
47084: gap of unknown length
48952: contig of 1868 bp in length
48953
49052: gap of unknown length
49053
49053
50861
50960: gap of 1808 bp in length
50961
52789: gap of unknown length
52889: gap of 1829 bp in length
52890
55277: gap of unknown length
55278
55377: contig of 2388 bp in length
55378
55378
56736: gap of unknown length
56737
56836: contig of 1359 bp in length
56837
56837
58697: gap of unknown length
58698
60311: contig of 1861 bp in length
60311
60312: contig of 1514 bp in length
60312
60411: gap of unknown length
60412
62535: contig of 2124 bp in length
62536
62536
62536: gap of unknown length
62536
64106: contig of 1471 bp in length
64206: gap of unknown length
64207
66041: contig of 1835 bp in length
66042
66141: gap of unknown length
66142
68585: contig of 2444 bp in length
68586
68586
68685: gap of unknown length
68686
71182: contig of 2437 bp in length
71183
71282: gap of unknown length
73823: contig of 2541 bp in length
73824
73923: gap of unknown length
73924
76163: contig of 2240 bp in length
76164
76263: gap of unknown length
76264
78665: contig of 2602 bp in length
78666
78865: gap of unknown length
78866
80953: contig of 1988 bp in length
80954
81053: gap of unknown length
81054
83204: contig of 2151 bp in length
83205
83304: gap of unknown length
83305
85471: contig of 2167 bp in length
85472
85571: gap of unknown length
88563: contig of 2992 bp in length
88564
88563: gap of unknown length
88564
91554: contig of 2881 bp in length
91555
91654: gap of unknown length
91655
94361: contig of 2707 bp in length
94461: gap of unknown length
94462
94462
99463: contig of 5002 bp in length

Query Match
Best Local Similarity 100.0%; Score 39; DB 2; Length 114558;
Pred. No. 9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 333 ACACGTGAATTCAGGCCATGAAGTTCGAAAGTAAG 371
Db 83890 ACACCTGAATTCAGGCCATGAAGTTCGAAAGTAAG 83852

RESULT 34
AC129612 160331 bp DNA linear HTG 24-AUG-2002
AC129612
LOCUS Rattus norvegicus clone CH230-129J12, *** SEQUENCING IN PROGRESS
DEFINITION
AC129612
AC129612
AC129612
AC129612.1 GI:22024295
```



KEYWORDS  
SOURCE  
ORGANISM

HTG; HTGS\_PHASE1  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 160331)  
Muzny,D.,Marle., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Alshrook,S., Amin,A., Anguiano,D.,  
Anylebech,V., Aoyagi,A., Ayodeji,A., Baca,E., Baden,H.,  
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
Bisvallo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
Byrnt,C., Bunay,C., Burch,C., Burrell,K., Calderon,E.,  
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
Davila,M.L., Davis,C., Davy-Carroll,L., De Andr.C., Dederich,D.,  
Delgado,O., Denison,S., Deramo,C., Ding,Y., Ditt,H., Divya,K.,  
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Gieria,W., Guevara,W.,  
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,  
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,  
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idubird,D., Jackson,A.,  
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
Kows,C., Kraft,C.L., Lebow,H., Levay,J., Lewis,L., Li,Z., Liu,J.,  
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
Lorenshew,L., Loulsegod,H., Lozado,R.J., Lu,K., Ma,J.,  
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milsavljecic,A.,  
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,  
Morris,S., Mundasa,M., Murphy,M., Nair,L., Nakervis,C., Neal,D.,  
Newton,N., Nguyen,N., Norris,S., Nwaokemele,C., Okwunonu,G.,  
Olanunnsagoun,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,  
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Polindexter,A.,  
Popovic,D., Primus,E., Pu,L., Puazo,M., Quatro,J., Racilin,E.,  
Reeves,K., Regier,M.A., Reigh,R., Rellay,B., Rellay,M., Ren,Y.,  
Reuter,M., Richards,S., Riggs,F., Rives,C., Rokeby,T., Rojas,A.,  
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saver,G., Scherer,S.,  
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shwartsbeyn,A.,  
Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sjoergren,E.,  
Song,X.-Z., Sorrelle,R., Sose,J., Steinle,M., Strong,R., Sutton,A.,  
Swatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,  
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Verh,V., Villaseana,D.,  
Walton,L., Walker,B., Wang,J., Wang,O., Wang,B., Warren,J.,  
Warren,R., Wei,X., White,F., Williams,G., Willison,R., Wlezyk,R.,  
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,  
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,  
Zhao,S., Dunn,D., von Niederhausen,A., Weiss,K., Smith,D.R.,  
Hol,R.A., Smith,H.O., Weinstock,G. and Gibbs,K.A.

TITLE  
JOURNAL

Unpublished  
2 (bases 1 to 160331)

REFERENCE  
AUTHORS  
JOURNAL

Direct Submission  
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 160331)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: KBGM  
Center clone name: CH230-129J12  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 92870 bases at least Q40  
Consensus quality: 98530 bases at least Q30  
Consensus quality: 102440 bases at least Q20  
-----

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 65 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1	1133:	contig of 1133 bp in length
*	1233:	gap of unknown length
*	1234	contig of 1137 bp in length
*	2271	gap of unknown length
*	2471	contig of 1028 bp in length
*	3459	gap of unknown length
*	3559	contig of 1015 bp in length
*	4514	gap of unknown length
*	4713:	gap of unknown length
*	4714	contig of 1054 bp in length
*	5767:	gap of unknown length
*	5768	gap of unknown length
*	5868	contig of 1484 bp in length
*	7351:	contig of 1484 bp in length
*	7352	gap of unknown length
*	7451:	gap of unknown length
*	7452	contig of 1632 bp in length
*	9083:	gap of unknown length
*	9084	contig of 1290 bp in length
*	10473:	gap of unknown length
*	10474	contig of 1607 bp in length
*	10574	contig of 1607 bp in length
*	12181	gap of unknown length
*	12281	contig of 1144 bp in length
*	13242:	gap of unknown length
*	13524:	gap of unknown length
*	13525	contig of 1526 bp in length
*	13521	gap of unknown length
*	15150:	gap of unknown length
*	16168:	contig of 1018 bp in length
*	16169	gap of unknown length
*	16268:	gap of unknown length
*	17904:	contig of 1636 bp in length
*	17905	gap of unknown length
*	18004:	gap of unknown length
*	18005	contig of 1030 bp in length
*	19035	gap of unknown length
*	19135	contig of 1079 bp in length
*	20213:	contig of 1079 bp in length
*	20214	gap of unknown length
*	20313:	gap of unknown length
*	22122:	contig of 1809 bp in length
*	22222:	gap of unknown length
*	22223	contig of 1937 bp in length
*	24159:	contig of 1937 bp in length
*	24259:	gap of unknown length
*	24260	contig of 1401 bp in length
*	25660:	contig of 1401 bp in length
*	25760:	gap of unknown length
*	25761	contig of 1342 bp in length
*	27102:	contig of 1342 bp in length
*	27203:	gap of unknown length
*	28639:	contig of 1437 bp in length
*	28640	gap of unknown length
*	28739:	gap of unknown length
*	29939:	contig of 1200 bp in length
*	29840	contig of 1200 bp in length
*	30039:	gap of unknown length
*	30040	contig of 1016 bp in length
*	31056	contig of 1016 bp in length
*	31056	gap of unknown length
*	32505:	contig of 1350 bp in length
*	32506	gap of unknown length
*	32606	contig of 1350 bp in length
*	33930:	contig of 1350 bp in length
*	33931	gap of unknown length
*	34031	contig of 1352 bp in length
*	35353	gap of unknown length
*	35452:	gap of unknown length
*	37015:	contig of 1563 bp in length

```

* 37016 37115: gap of unknown length
* 37116 39630: contig of 2515 bp in length
* 39630 39730: gap of unknown length
* 39730 41572: contig of 1842 bp in length
* 41572 41672: gap of unknown length
* 41672 43363: contig of 1691 bp in length
* 43363 43463: gap of unknown length
* 43463 46153: contig of 2690 bp in length
* 46153 46253: gap of unknown length
* 46253 47753: contig of 1500 bp in length
* 47753 47853: gap of unknown length
* 47853 49651: contig of 1798 bp in length
* 49651 49751: gap of unknown length
* 49751 51191: contig of 1440 bp in length
* 51191 51291: gap of unknown length
* 51291 52673: contig of 1382 bp in length
* 52673 52773: gap of unknown length
* 52773 54972: contig of 2199 bp in length
* 54972 55072: gap of unknown length
* 55072 57153: contig of 2081 bp in length
* 57153 57253: gap of unknown length
* 57253 59614: contig of 2361 bp in length
* 59614 59714: gap of unknown length
* 59714 61186: contig of 1472 bp in length
* 61186 61286: gap of unknown length
* 61286 63329: contig of 2043 bp in length
* 63329 63429: gap of unknown length
* 63429 64740: contig of 1311 bp in length
* 64740 64841: gap of unknown length
* 64841 66353: contig of 1513 bp in length
* 66353 66453: gap of unknown length
* 66453 68773: contig of 2320 bp in length
* 68773 68874: gap of unknown length
* 68874 72002: contig of 3129 bp in length
* 72002 72102: gap of unknown length
* 72102 74179: contig of 2077 bp in length
* 74179 74279: gap of unknown length
* 74279 76076: contig of 1797 bp in length
* 76076 76176: gap of unknown length
* 76176 77703: contig of 1527 bp in length
* 77703 77803: gap of unknown length
* 77803 80423: contig of 2620 bp in length
* 80423 80524: gap of unknown length
* 80524 84110: contig of 3587 bp in length
* 84110 84210: gap of unknown length
* 84210 87326: contig of 3116 bp in length
* 87326 87426: gap of unknown length
* 87426 89663: contig of 2437 bp in length
* 89663 89864: gap of unknown length

```

Query Match 2.4%; Score 39; DB 2; Length 160331;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-11;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 333 ACACCTGATTCAGGCCAATGAGTTGCGAAGTGAAG 371
      |||||||
Db 156808 ACACCTGATTCAGGCCAATGAGTTGCGAAGTGAAG 156846

RESULT 35
MM089404
LOCUS MM089404 246 bp mRNA linear ROD 23-SEP-1998
DEFINITION Mus musculus strain BALB/c inosine-5'-monophosphate dehydrogenase 2
      mRNA, partial cds.
ACCESSION U89404.1 GI:3642640
VERSION U89404.1
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 246)
AUTHORS Chu,C.C. and Paul,W.E.

```

```

TITLE Expressed genes in interleukin-4 treated B cells identified by cDNA
JOURNAL Mol. Immunol. 35 (8), 487-502 (1998)
MEDLINE 99012997
PUBMED 9798653
REFERENCE 2 (bases 1 to 246)
AUTHORS Chu,C.C. and Paul,W.E.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1997) Lab. of Immunology, NIAID, Building 10,
      Room 11N311, Bethesda, MD 20892-1892, USA
FEATURES
      source
      1..246
      /organism="Mus musculus"
      /strain="BALB/c"
      /db_xref="taxon:10090"
      /cell_type="B cells treated with LPS and IL-4 for 12 and
      36 hours"
      /tissue_type="spleen"
      <1..>246
      /note="rate limiting enzyme in guanine nucleotide
      biosynthesis; IMP dehydrogenase"
      /codon_start=1
      /product="inosine-5'-monophosphate dehydrogenase 2"
      /protein_id="AAC36511.1"
      /db_xref="GI:3642641"
      /translation="GGNVYTAQAKNLIDAGVDALRVGSGSICITQEVILACGRQA
      TAVYVSEPARFCFVPIADGCIQVNGHAKALAGAP"
BASE COUNT 51 a 57 c 75 g 63 t
ORIGIN

```

Query Match 2.3%; Score 38; DB 10; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1131 ATTGCTGATGAGAGATCCAAATGCGTCATATTC 1168
      |||||||
Db 184 ATTGCTGATGAGAGATCCAAATGCGTCATATTC 221

```

```

RESULT 36
AC103220
LOCUS AC103220 168808 bp DNA linear HTG 13-JUL-2002
DEFINITION Rattus norvegicus clone CH230-48M6, *** SEQUENCING IN PROGRESS ***
ACCESSION AC103220
VERSION AC103220.3 GI:21731183
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 168808)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
      Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T.,
      Barbarta,J., Benton,J., Binaghi,K., Blankenburg,K., Bonini,D.,
      Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
      Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
      Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
      Chen,G., Chen,R., Chen,Z., Chowdhury,T., Christopoulos,C.,
      Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
      Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
      Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
      Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
      Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
      Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
      Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
      Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
      Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J.,
      Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
      Homsi,E., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
      Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,
      Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

```

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,  
Lorodo,R.J., Lu,X., Lucier,A., Lucier,R., Lunde,R., Ma,J.,  
Masreshari,M., Mapua,P., Martin,R., Martindale,A., Mattheze,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunnu,G.,  
Oragunne,N., Oviedo,R., Pace,A., Payton,B., Perry,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pul-L., Oulles,M., Ren,Y.,  
Rivers,M., Rojas,A., RojudoKan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonaikhe,T., Sparks,A., Stanley,F., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

-----  
Direct Submission  
Unpublished  
2 (bases 1 to 168808)  
Worley,K.C.  
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 168808)  
Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:17974711.

-----  
Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
-----  
Project Information  
Center project name: GJMO  
Center clone name: CH230-48M6  
-----  
Summary Statistics  
Sequencing vector: plasmid:  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 125544 bases at least Q40  
Consensus quality: 130223 bases at least Q80  
Consensus quality: 133993 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 63 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1360 1459: contig of 1359 bp in length  
1460 2564: gap of unknown length  
2565 2664: gap of unknown length  
2665 4113: contig of 1449 bp in length  
4114 4214: gap of unknown length  
4214 5667: contig of 1454 bp in length  
5667 5768: gap of unknown length  
5768 7479: contig of 1712 bp in length  
7479 7580: gap of unknown length  
7580 8718: contig of 1139 bp in length  
8718 8819: gap of unknown length

8819 10257: contig of 1439 bp in length  
10258 10357: gap of unknown length  
10358 11631: contig of 1274 bp in length  
11632 11731: gap of unknown length  
11732 13108: contig of 1377 bp in length  
13109 13208: gap of unknown length  
13209 14306: contig of 1098 bp in length  
14307 14407: gap of unknown length  
14407 14527: contig of 1021 bp in length  
14528 15577: gap of unknown length  
15578 16648: contig of 1121 bp in length  
16649 16748: gap of unknown length  
16749 19101: contig of 2353 bp in length  
19102 21321: gap of unknown length  
21322 21421: contig of 2120 bp in length  
21422 23074: gap of unknown length  
23074 23174: contig of 1653 bp in length  
23175 24513: contig of 1339 bp in length  
24514 24613: gap of unknown length  
24614 25820: contig of 1207 bp in length  
25821 27102: contig of 1182 bp in length  
27103 27202: gap of unknown length  
27203 28457: contig of 1235 bp in length  
28458 28597: gap of unknown length  
28598 30228: contig of 1631 bp in length  
30229 31341: gap of unknown length  
31342 31441: contig of 1013 bp in length  
31442 33385: gap of unknown length  
33386 33495: gap of unknown length  
33496 34901: contig of 1406 bp in length  
34902 35001: gap of unknown length  
35002 37643: contig of 2642 bp in length  
37644 37744: gap of unknown length  
37744 39049: contig of 1306 bp in length  
39050 39149: gap of unknown length  
39150 40518: contig of 1369 bp in length  
40519 40618: gap of unknown length  
40619 42516: contig of 1898 bp in length  
42517 42616: gap of unknown length  
42617 44169: contig of 1553 bp in length  
44170 44269: gap of unknown length  
44270 45701: contig of 1432 bp in length  
45702 45801: gap of unknown length  
45802 47590: contig of 1789 bp in length  
47591 47690: gap of unknown length  
47691 48832: contig of 1142 bp in length  
48833 48932: gap of unknown length  
48933 51270: contig of 2338 bp in length  
51271 51370: gap of unknown length  
51371 53607: contig of 2237 bp in length  
53608 53707: gap of unknown length  
53708 56265: contig of 2558 bp in length  
56266 56365: gap of unknown length  
56366 59684: contig of 3319 bp in length  
59685 59784: gap of unknown length  
59785 61687: contig of 1903 bp in length  
61688 61787: gap of unknown length  
61788 63750: contig of 1963 bp in length  
63751 63850: gap of unknown length  
63851 65690: contig of 1840 bp in length  
65691 65790: gap of unknown length  
65791 69754: contig of 3964 bp in length  
69755 69854: gap of unknown length  
69855 71431: contig of 1377 bp in length  
71432 71531: gap of unknown length  
71532 74389: contig of 2858 bp in length  
74390 74489: gap of unknown length  
74490 77318: contig of 2829 bp in length  
77319 80598: gap of unknown length  
80598 8819: contig of 3180 bp in length

```

      * 80599      80698: gap of unknown length
      * 80699      83492: contig of 2794 bp in length
      * 83493      83592: gap of unknown length
      * 83593      87803: contig of 4211 bp in length
      * 87804      87903: gap of unknown length
      * 87904      91190: contig of 3287 bp in length
      * 91191      91290: gap of unknown length
      * 91291      94503: contig of 3213 bp in length
      * 94504      94603: gap of unknown length
      * 94604      98264: contig of 3661 bp in length
      * 98265      98364: gap of unknown length
      * 98365      102082: contig of 3718 bp in length
      * 102083      102182: gap of unknown length
      * 102183      107090: contig of 4908 bp in length
      * 107091      107190: gap of unknown length
      * 107191      111367: contig of 4077 bp in length
      * 111368      111367: gap of unknown length
      * 111368      115934: contig of 4567 bp in length
      * 115935      116034: gap of unknown length
      * 116035      119693: contig of 3659 bp in length
      * 119694      119793: gap of unknown length
      * 119794      125196: contig of 5403 bp in length

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 168808;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1335 TATTTCAGTGAAGCTGACAAATCAATCAAGTGGCC 1367
Db 75466 TATTTCAGTGAAGCTGACAAATCAATCAAGTGGCC 75498

RESULT 37
AR119642
LOCUS AR119642 44 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6153398.
ACCESSION AR119642
VERSION AR119642.1 GI:14102341
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 44)
AUTHORS Collart,F.R. and Huberman,E.
TITLE Method to identify specific inhibitors of IMP dehydrogenase
JOURNAL Patent: US 6153398-A 1 28-NOV-2000;
FEATURES
Location/Qualifiers
source
1..44
BASE COUNT 11 a 10 c 13 g 10 t
ORIGIN
11 a 10 c 13 g 10 t

Query Match
Best Local Similarity 1.8%; Score 30; DB 6; Length 44;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 48 ATGGCCGACTACTGATGATGAGGGGACG 77
Db 15 ATGGCCGACTACTGATGATGAGGGGACG 44

RESULT 38
HUMYO19A04
LOCUS HUMYO19A04 687 bp mRNA linear PRI 04-AUG-1998
DEFINITION Homo sapiens full length insert cDNA Y019A04.
ACCESSION AF075074
VERSION AF075074.1 GI:3377615
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 687)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

AUTHORS Woessner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,
March,G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theising,B.,
Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
Gibbons M., Harvey,N., Pape,D., Chamberlain,A., Morales,R.,
Schurk,R., Ritter,E., Kohn,S., Skaller,T., Beynemer,K., Hiller,L.,
Wilson,R. and Waterson,R.
TITLE Full Clone Sequencing of the Longest Available Member from Each
Unpublished
JOURNAL 2 (bases 1 to 687)
REFERENCE
AUTHORS Waterson,R.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
SUBMITTED BY:
Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:est@watson.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. NO
attempt has been made to verify whether this corresponds to the
full-length of the original mRNA from which it was derived. We
have tried to obtain double-stranded, or double chemistry sequence
across the entire clone, but potentially, there are areas in the
sequence where this level of coverage was not achieved.
Nevertheless, we are confident of the accuracy of this sequence as
all regions of low quality, as defined by PHRAP (P. Green, in
preparation), were visually inspected and edited accordingly. The
consensus quality values for this sequence have been submitted
separately.

The location of this clone is unknown.
FEATURES
Location/Qualifiers
source
1..687
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="274158"
/clone_lib="Soares_fetal_liver_spleen_1NFS"
repeat_region
402..687
/rpt_family="ALU"
BASE COUNT 195 a 178 c 165 g 149 t
ORIGIN

Query Match
Best Local Similarity 1.8%; Score 29; DB 9; Length 687;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1617 GTTTTTCATATAAAGTTTGAAGA 1645
Db 2 GTTTTTCATATAAAGTTTGAAGA 30

RESULT 39
AX348472
LOCUS AX348472 6193 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 167 from Patent WO0202806.
ACCESSION AX348472
VERSION AX348472.1 GI:18614508
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL Patent: WO 0202806-A 167 10-JAN-2002;
FEATURES
Epigenomics AG (DE)
Location/Qualifiers
source
1..6193
/organism="synthetic construct"

```

/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1404 a 124 c 1628 g 3037 t

ORIGIN

Query Match 1.7%; Score 28; DB 6; Length 6193;  
Best Local Similarity 100.0%; Pred. No. 0.00039;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 AACGATCGCTGGGATGTTTGGAGC 444  
|||||

Db 2561 AACGATCGCTGGGATGTTTGGAGC 2588

RESULT 40  
AC107280

LOCUS AC107280 114558 bp DNA linear HTG 13-JUL-2002  
DEFINITION Rattus norvegicus clone CH230-81H4. \*\*\* SEQUENCING IN PROGRESS \*\*\*  
58 unordered pieces.  
AC107280  
VERSION AC107280.3 GI:21737470  
KEYWORDS HTG: HTGS PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 114558)  
Muzny,D.M., Adams,C., Adio-oduola,B., All-ouma,F.R., Allen,C.,  
Alstbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayde,M., Banks,T.,  
Barbieri,J., Benton,J., Bimge,K., Blankenburg,K., Bonin,D.,  
Bouay,J., Bowie,S., Brleya,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carson,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,  
Doutwalte,K.J., Draper,H., Dugan-Rocha,S., Dublin,K.J.,  
Earmhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,B., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoen,E., Kelly,S., Khan,U., King,L., Korva,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
Miner,G., Minter,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,S., Noun,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenko,S., Ogih,M., Okunogu,G.,  
Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peary,J., Perez,L.,  
Peters,L., Pichens,R., Primus,E., Pu,L.L., Qureshi,M., Ren,Y.,  
Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Severy,G.,  
Scheider,S., Scott,G., Shen,H., Shoshchari,N., Sisson,I.,  
Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansy,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vison,R., Wang,O.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,S., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,D.,  
Weinstock,G., and Gibbs,R.

TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 114558)  
AUTHORS Worley,K.C.  
TITLE Direct Submission

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Submitted (18-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 114558)  
Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18702062.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPMW  
Center clone name: CH230-81H4  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 39289 bases at least Q40  
Consensus quality: 62371 bases at least Q30  
Consensus quality: 65704 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank-draft-data.html>).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 58 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1407: contig of 1407 bp in length  
1408 1507: gap of unknown length  
1508 2785: contig of 1278 bp in length  
2786 2885: gap of unknown length  
2886 4306: contig of 1421 bp in length  
4307 4407: gap of unknown length  
4407 5791: contig of 1385 bp in length  
5792 5891: gap of unknown length  
5892 7219: contig of 1328 bp in length  
7220 7319: gap of unknown length  
7320 8441: contig of 1122 bp in length  
8442 8541: gap of unknown length  
8542 9964: contig of 1422 bp in length  
9964 10064: gap of unknown length  
10064 11368: contig of 1305 bp in length  
11368 11469: gap of unknown length  
11469 12924: contig of 1456 bp in length  
12925 13025: gap of unknown length  
13025 14075: contig of 1050 bp in length  
14075 14175: gap of unknown length  
14175 15767: contig of 1593 bp in length  
15768 15867: gap of unknown length  
15868 17131: contig of 1264 bp in length  
17132 17231: gap of unknown length  
17232 18519: contig of 1288 bp in length  
18520 18620: gap of unknown length  
18620 19671: contig of 1053 bp in length  
19671 19771: gap of unknown length  
19772 21140: contig of 1369 bp in length  
21141 21240: gap of unknown length  
21241 22449: contig of 1209 bp in length  
22450 23904: gap of unknown length  
23905 24004: gap of unknown length  
24005 25150: contig of 1146 bp in length  
25151 25250: gap of unknown length  
25250 26269: contig of 1019 bp in length

```
*
* 26270 26369: gap of unknown length
* 26370 27985: contig of 1616 bp in length
* 27986 28085: gap of unknown length
* 28086 29407: contig of 1322 bp in length
* 29408 29507: gap of unknown length
* 29508 30634: contig of 1127 bp in length
* 30635 30734: gap of unknown length
* 30735 31927: contig of 1193 bp in length
* 31928 32027: gap of unknown length
* 32028 33335: contig of 1308 bp in length
* 33336 33435: gap of unknown length
* 33436 35111: contig of 2076 bp in length
* 35112 35611: gap of unknown length
* 35612 37208: contig of 1597 bp in length
* 37209 37308: gap of unknown length
* 37309 38952: contig of 1644 bp in length
* 38953 39052: gap of unknown length
* 39053 40524: contig of 1472 bp in length
* 40525 40624: gap of unknown length
* 40625 42194: contig of 1570 bp in length
* 42195 42294: gap of unknown length
* 42295 43782: contig of 1488 bp in length
* 43783 43882: gap of unknown length
* 43883 45446: contig of 1564 bp in length
* 45447 46984: contig of 1438 bp in length
* 46985 47084: gap of unknown length
* 47085 48952: contig of 1868 bp in length
* 48953 49052: gap of unknown length
* 49053 50860: contig of 1808 bp in length
* 50861 50960: gap of unknown length
* 50961 52790: contig of 1829 bp in length
* 52790 55277: gap of unknown length
* 55277 55378: contig of 2388 bp in length
* 55378 55378: gap of unknown length
* 55378 56736: gap of 1359 bp in length
* 56737 56836: gap of unknown length
* 56837 58697: contig of 1861 bp in length
* 58698 58797: gap of unknown length
* 58798 60311: contig of 1514 bp in length
* 60312 60411: gap of unknown length
* 60412 62535: contig of 2124 bp in length
* 62535 62636: gap of unknown length
* 62636 64106: contig of 1471 bp in length
* 64107 64206: gap of unknown length
* 64207 66041: contig of 1835 bp in length
* 66042 66141: gap of unknown length
* 66142 68585: contig of 2444 bp in length
* 68586 68685: gap of unknown length
* 68686 71182: contig of 2497 bp in length
* 71183 71282: gap of unknown length
* 71283 73823: contig of 2541 bp in length
* 73824 73923: gap of unknown length
* 73924 76163: contig of 2240 bp in length
* 76164 76263: gap of unknown length
* 76264 78865: contig of 2602 bp in length
* 78866 78965: gap of unknown length
* 78966 80953: contig of 1988 bp in length
* 80954 81053: gap of unknown length
* 81054 83204: contig of 2151 bp in length
* 83205 83304: gap of unknown length
* 83305 85471: contig of 2167 bp in length
* 85472 85571: gap of unknown length
* 85572 88563: contig of 2992 bp in length
* 88564 88664: gap of unknown length
* 88664 91554: contig of 2891 bp in length
* 91555 91654: gap of unknown length
* 91655 94361: contig of 2707 bp in length
* 94362 94461: gap of unknown length
* 94462 99463: contig of 5002 bp in length
*
```

Query Match 1.7%; Score 28; DB 2; Length 114558;  
Best Local Similarity 100.0%; Pred. No. 0.00031;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1216 TGGCTGCCACCTGAGAGCCCTGTGTGA 1243  
|||||  
Db 49648 TGGCTGCCACCTGAGAGCCCTGTGTGA 49675

RESULT 41  
AR119643/c  
LOCUS AR119643 42 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 2 from patent US 6153398.  
ACCESSION AR119643  
VERSION AR119643.1 GI:14102342  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 42)  
AUTHORS Collart,F.R. and Huberman,E.  
TITLE Method to identify specific inhibitors of IMP dehydrogenase  
JOURNAL Patent: US 6153398-A 2 28-NOV-2000;  
FEATURES  
source 1..42  
BASE COUNT 15 a 10 c 7 g 10 t  
ORIGIN

Query Match 1.6%; Score 26; DB 6; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.0084;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1568 TTGCTATGAGACGCGCTTTCTGAA 1593  
|||||  
Db 42 TTGCTATGAGACGCGCTTTCTGAA 17

RESULT 42  
AX190187 50 bp DNA linear PAT 08-AUG-2001  
LOCUS AX190187/c  
DEFINITION Sequence 366 from Patent WO0147942.  
ACCESSION AX190187  
VERSION AX190187.1 GI:15143563  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Shinkets,R.A. and Leach,M.  
TITLE Nucleic acids containing single nucleotide polymorphisms and  
methods of use thereof  
JOURNAL Patent: WO 0147942-A 366 05-JUL-2001;  
Curagen Corporation (US)  
FEATURES  
source Location/Qualifiers  
1..50  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="2 of 2 allelic variants (365 is other entry)"  
misc\_feature 25..26  
/note="Nucleotide deleted between bases 25 and 26"  
Accession number C943922096"

BASE COUNT 12 a 20 c 12 g 6 t  
ORIGIN

Query Match 1.6%; Score 26; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 27 CGCGGGTGTCGTGTGGCCATGGC 52  
|||||  
Db 50 CGCGGGTGTCGTGTGGCCATGGC 25

```
RESULT 43
LOCUS AX190185/C 51 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 364 from Patent WO0147942.
ACCESSION AX190185
VERSION AX190185.1 GI:15143561
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Shinkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0147942-A 364 05-JUL-2001;
Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="1 of 2 allelic variants (364 is other
entry)-Accession number C943922096"
BASE COUNT 6 a 18 c 12 g 15 t
ORIGIN
Query Match 1.6%; Score 26; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 AGAGGAGACCTGTGTAGCCCC 614
Db 51 AGAGGAGACCTGTGTAGCCCC 26

RESULT 44
LOCUS AX418043/C 35 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 53 from Patent WO0185952.
ACCESSION AX418043
VERSION AX418043.1 GI:21523088
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Krystek,S.R., Sheriff,S., Wilmer,M.R., Hollenbaugh,D.L., Yan,N.,
Moutavieff,J.E., Einspahr,H.M. and Kish,K.
TITLE Modified inosine 5'-monophosphate dehydrogenase polypeptides and
uses thereof
JOURNAL Patent: WO 0185952-A 53 15-NOV-2001;
Bristol-Myers Squibb Co. (US)
FEATURES
source location/Qualifiers
1..35
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 10 a 9 c 7 g 9 t
ORIGIN
Query Match 1.5%; Score 25; DB 6; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1570 CGATGAGAGCGGCTTTCTGAAA 1594
Db 35 CGATGAGAGCGGCTTTCTGAAA 11

RESULT 45
LOCUS AX190184/C 51 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 363 from Patent WO0147942.
```

```
ACCESSION AX190184
VERSION AX190184.1 GI:15143560
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Shinkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: WO 0147942-A 363 05-JUL-2001;
Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="1 of 2 allelic variants (364 is other
entry)-Accession number C943922096"
BASE COUNT 7 a 18 c 11 g 15 t
ORIGIN
Query Match 1.5%; Score 25; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 AGAGGAGACCTGTGTAGCCCC 613
Db 51 AGAGGAGACCTGTGTAGCCCC 27

RESULT 46
LOCUS AX418042 33 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 52 from Patent WO0185952.
ACCESSION AX418042
VERSION AX418042.1 GI:21523087
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Krystek,S.R., Sheriff,S., Wilmer,M.R., Hollenbaugh,D.L., Yan,N.,
Moutavieff,J.E., Einspahr,H.M. and Kish,K.
TITLE Modified inosine 5'-monophosphate dehydrogenase polypeptides and
uses thereof
JOURNAL Patent: WO 0185952-A 52 15-NOV-2001;
Bristol-Myers Squibb Co. (US)
FEATURES
source location/Qualifiers
1..33
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 7 a 8 c 9 g 9 t
ORIGIN
Query Match 1.5%; Score 24; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ATGGCCGACTACCTGATAGTGG 71
Db 10 ATGGCCGACTACCTGATAGTGG 33

RESULT 47
LOCUS AX418045/C 30 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 55 from Patent WO0185952.
ACCESSION AX418045
VERSION AX418045.1 GI:21523090
KEYWORDS
SOURCE human.
```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
  1
AUTHORS Kravets,S.R., Shierliff,S., Wimer,M.R., Hollenbaugh,D.L., Yan,N.,
Mouravieff,J.E., Einspahr,H.M. and Kish,K.
TITLE Modified inosine 5'-monophosphate dehydrogenase polypeptides and
uses thereof
JOURNAL Patent: WO 0185952-A 55 15-NOV-2001;
Bristol-Myers Squibb Co. (US)
FEATURES
  source
    1..30
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT      5 a      8 c      4 g      13 t
ORIGIN
Query Match      1.4%; Score 23; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 GTTCGGAAGTGAAGAATATGA 379
      |||||||||||||||||||
Db 30 GTTCGGAAGTGAAGAATATGA 8

RESULT 48
ARI41891
LOCUS ARI41891 344 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6147194.
ACCESSION ARI41891
VERSION ARI41891.1 GI:15101407
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 344)
AUTHORS Collart,F.R. and Huberman,E.
TITLE Eukaryotic IMPDH polynucleotide and antibody compositions
JOURNAL Patent: US 6147194-A 3 14-NOV-2000;
FEATURES
  source
    1..344
    /organism="unknown"
BASE COUNT      93 a      76 c      101 g      74 t
ORIGIN
Query Match      1.4%; Score 23; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 GCCATCATTTGCCGACAGACCT 727
      |||||||||||||||||||
Db 295 GCCATCATTTGCCGACAGACCT 317

RESULT 49
ARI41892/C
LOCUS ARI41892 344 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 4 from patent US 6147194.
ACCESSION ARI41892
VERSION ARI41892.1 GI:15101408
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 344)
AUTHORS Collart,F.R. and Huberman,E.
TITLE Eukaryotic IMPDH polynucleotide and antibody compositions
JOURNAL Patent: US 6147194-A 4 14-NOV-2000;
FEATURES
  source
    1..344
    /organism="unknown"
BASE COUNT      74 a      101 c      76 g      93 t

```

```

ORIGIN
Query Match      1.4%; Score 23; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 GCCATCATTTGCCGACAGACCT 727
      |||||||||||||||||||
Db 50 GCCATCATTTGCCGACAGACCT 28

RESULT 50
I06573
LOCUS I06573 344 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 4 from Patent WO 9001545.
ACCESSION I06573
VERSION I06573.1 GI:589655
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 344)
AUTHORS Collart,F.R. and Huberman,E.
TITLE METHOD AND MATERIALS RELATING TO IMPDH AND GMP PRODUCTION
JOURNAL Patent: WO 9001545-A 4 22-FEB-1990;
FEATURES
  source
    1..344
    /organism="unknown"
BASE COUNT      93 a      76 c      101 g      74 t
ORIGIN
Query Match      1.4%; Score 23; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 GCCATCATTTGCCGACAGACCT 727
      |||||||||||||||||||
Db 295 GCCATCATTTGCCGACAGACCT 317

```

Search completed: February 13, 2003, 01:05:35  
 Job time : 4980 secs







```
RESULT 2
ABV29489
ID ABV29489 standard: cDNA: 2277 BP.
XX
AC ABV29489:
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29480.
XX
KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
XX
XX pharmacogenomic marker; gene: SS.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-0505171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI: 2001-662795/76.
XX
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1: Page 6318: 11750bp: English.
XX
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (1) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 2277 BP: 644 A: 568 C: 653 G: 412 T: 0 other
SQ

Query Match      68.1% Score 1127: DB 23: Len 2277:
Best Local Similarity 99.4% Pred. No. 0:
Matches 1627: Conservative 0: Mismatches 10: Indels 0: Gaps 0:
```

```
QY 189 CAGGTGACCTGACTTCTGTCTGACCAGAAATACCTTTAAGACCCCACTGTTTCC 248
    |||||||
DB 273 CAGGTGACCTGACTTCTGTCTGACCAGAAATACCTTTAAGACCCCACTGTTTCC 332
QY 249 TCTCCATGAGACAGTACAGAGGCTGGATGGCCATGACATGGCGCTTACAGCGGT 308
    |||||||
DB 333 TCTCCATGAGACAGTACAGAGGCTGGATGGCCATGACATGGCGCTTACAGCGGT 392
QY 309 ATTGGCTTCATCCACCAACCTGACACCTGATTCCTCAGGCCATGACATGATGGAAGTG 368
    |||||||
DB 393 ATTGGCTTCATCCACCAACCTGACACCTGATTCCTCAGGCCATGACATGATGGAAGTG 452
QY 369 AGAATATGAGACAGGATTCATCAGAGACCTGTGTCTCAGCCCAAGATTCCTGTG 428
    |||||||
DB 453 AGAATATGAGACAGGATTCATCAGAGACCTGTGTCTCAGCCCAAGATTCCTGTG 512
QY 429 CGGGATGTTTTGAGGCGAAGGCCCGGATGTTCTCGGGTATCCCAATCAGACACACA 488
    |||||||
DB 513 CGGGATGTTTTGAGGCGAAGGCCCGGATGTTCTCGGGTATCCCAATCAGACACACA 572
QY 489 GGGCGATGGGAGCGCGTGTGGCATCATCTCCAGGACATTTGTTCTCAA 548
    |||||||
DB 573 GGGCGATGGGAGCGCGTGTGGCATCATCTCCAGGACATTTGTTCTCAA 632
QY 549 GAGGAGAACATGACGTGTTCTTGAAGATATGACAAAGGAGGAGACTTGGTGTA 608
    |||||||
DB 633 GAGGAGAACATGACGTGTTCTTGAAGATATGACAAAGGAGGAGACTTGGTGTA 692
QY 609 GCCCCCCGACATCACACTGAAGAGGCAATGAATTCGTGAGGCGACAGAAAGGA 668
    |||||||
DB 693 GCCCTTCGAGGATCACACAGGAAGGAGCAATGAATTCGTGAGGCGACAGAAAGGA 752
QY 669 AAGTGGCCATTTAATGATGATGATGAGCTGTGGCCATCTTCCCGGACAGACCTG 728
    |||||||
DB 753 AAGTGGCCATTTAATGATGATGATGAGCTGTGGCCATCTTCCCGGACAGACCTG 812
QY 729 AAGAAGATCGGAGCTAACCCACTAGCCTCCAAAGATGCCAAGAACAGCTGTGTGG 788
    |||||||
DB 813 AAGAAGATCGGAGCTAACCCACTAGCCTCCAAAGATGCCAAGAACAGCTGTGTGG 872
QY 789 GCAGCCATTCGACCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 848
    |||||||
DB 873 GCAGCCATTCGACCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 932
QY 849 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
    |||||||
DB 933 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 992
QY 909 AAGTACATCAAAAGCAAAATACCTTATTCGATGATGATGATGATGATGATGATGATG 968
    |||||||
DB 993 AAGTACATCAAAAGCAAAATACCTTATTCGATGATGATGATGATGATGATGATGATG 1052
QY 969 GCCCAGGCAAGAACCTATTGATGAGGATGATGATGATGATGATGATGATGATGATG 1028
    |||||||
DB 1053 GCCCAGGCAAGAACCTATTGATGAGGATGATGATGATGATGATGATGATGATGATG 1112
QY 1029 GGGTCCATTCGATTCATTCAGGAGAGTGTGGCTGTGGGCGGCCCAAGCAATGAGAGTG 1088
    |||||||
DB 1113 GGGTCCATTCGATTCATTCAGGAGAGTGTGGCTGTGGGCGGCCCAAGCAATGAGAGTG 1172
QY 1089 TACAGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
    |||||||
DB 1173 TACAGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1232
QY 1149 CAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
    |||||||
DB 1233 CAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1292
QY 1209 TCTCTCTGAGCTGCAACACTGAGGCCCTGTGTAATCTTTTTCGATGGATTCGG 1268
    |||||||
DB 1293 TCTCTCTGAGCTGCAACACTGAGGCCCTGTGTAATCTTTTTCGATGGATTCGG 1352
```

```
OY 1269 CTTAAGAATATTCGCGGTATGGGTTCTCTGATGCCATGACAGCACCCTCAGACGCCAG 1328
DB 1353 CTTAAGAAATATTCGCGGTATGGGTTCTCTGATGCCATGACAGCACCCTCAGACGCCAG 1412
OY 1329 AACGATATTTTCAGTGAAGCTGCAAAATCAAGTGGCCCGAGGAGTGTCTGTGTG 1388
DB 1413 AACGATATTTTCAGTGAAGCTGCAAAATCAAGTGGCCCGAGGAGTGTCTGTGTG 1472
OY 1389 CAGGACAAAGGGGCATCCACAATTTGTCCTTACCTGATTCGTCGATCCCAACACTCA 1448
DB 1473 CAGGACAAAGGGGCATCCACAATTTGTCCTTACCTGATTCGTCGATCCCAACACTCA 1532
OY 1449 TGCACGACATTTGTCGCCAAGACCTTGACCAAGTCCGAGCCATGATGTACTCTGGGGAG 1508
DB 1533 TGCACGACATTTGTCGCCAAGACCTTGACCAAGTCCGAGCCATGATGTACTCTGGGGAG 1592
OY 1509 CTTTAACTTTGAGAAGAAAGCTCTTCAGCCCGCAGAGTGGAGTGGCTCCATAGCCTCCAT 1568
DB 1593 CTTTAACTTTGAGAAGAAAGCTCTTCAGCCCGCAGAGTGGAGTGGCTCCATAGCCTCCAT 1652
OY 1569 TCGTATGAGAAGGGGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCA 1628
DB 1653 TCGTATGAGAAGGGGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCA 1712
OY 1629 ATTAAAGTTTAGAAGA 1645
DB 1713 ATTAAAGTTTAGAAGA 1729

RESULT 3
ABV23631
ID ABV23631 standard; cDNA; 1677 BP.
XX
AC ABV23631;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23622.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI: 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 4329; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
```

```
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 1677 BP; 428 A; 446 C; 445 G; 358 T; 0 other;
XX
Query Match 55.3%; Score 914; DB 23; Length 1677;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1364; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OY 187 ACCAGGTGACCTGACTTGTGCTGTGACCAAGAAATCACTTTAGACCCCACTGTTT 246
DB 132 ACCAGGTGACCTGACTTGTGCTGTGACCAAGAAATCACTTTAGACCCCACTGTTT 191
OY 247 CCTCTCCATGGACACAGTACAGAGGCTGGATGGCCATAGCAATGGCTTACAGGCG 306
DB 192 CCTCTCCATGGACACAGTACAGAGGCTGGATGGCCATAGCAATGGCTTACAGGCG 251
OY 307 GTATTGGCTTCATCCCACTGACACACTGTAATTTCCAGCCCAATGAATTCGGAAG 366
DB 252 GTATTGGCTTCATCCCACTGACACACTGTAATTTCCAGGCCCAATGAATTCGGAAG 311
OY 367 TGAAGAAATATGAAAGGAGTTCATCAGACACCTTGTGCTCAGACCCCAAGATCGCG 426
DB 312 TGAAGAAATATGAAAGGAGTTCATCAGACACCTTGTGCTCAGACCCCAAGATCGCG 371
OY 427 TCGGGATGTTTTTGAAGCCCAAGGCCCGCATGTTTCTGCGGTATCCCAATCAGACA 486
DB 372 TCGGGATGTTTTTGAAGCCCAAGGCCCGCATGTTTCTGCGGTATCCCAATCAGACA 431
OY 487 CAGGCGGATGGGAGCGGCTGTGGGATCATCTCCGAGGAGCATTTGATTTCCTCA 546
DB 432 CAGGCGGATGGGAGCGGCTGTGGGATCATCTCCGAGGAGCATTTGATTTCCTCA 491
OY 547 AAGAGGAGAACATGACTGTTTCTTGAAGAGATTAATGACAAAGAGGAAAGACTTGGTGG 606
DB 492 AAGAGGAGAACATGACTGTTTCTTGAAGAGATTAATGACAAAGAGGAAAGACTTGGTGG 551
OY 607 TAGCCCCCGCAGCATCAGACTGAAGAGGCAATGAATTTCTGACGCCAGCAGAGAGG 666
DB 552 TAGCCCCCGCAGCATCAGACTGAAGAGGCAATGAATTTCTGACGCCAGCAGAGAGG 611
OY 667 GAAAGTTGCCCATTTGAAATGAAGATGATGAGCTTGTGGCCATGTTGCCGAGACGACC 726
DB 612 GAAAGTTGCCCATTTGAAATGAAGATGATGAGCTTGTGGCCATGTTGCCGAGACGACC 671
OY 727 TGAAGAGAAATGGGAGTACCTAGCCTCCAAAATGCCAAGAAAGAGTGTGTGTG 786
DB 672 TGAAGAGAAATGGGAGTACCTAGCCTCCAAAATGCCAAGAAAGAGTGTGTGTG 731
OY 787 GGGCAGCATTGGGACTCATGAGAGTACAGAGTATAGGCTGGACTGTCGCCAGGCTG 846
DB 732 GGGCAGCATTGGGACTCATGAGAGTACAGAGTATAGGCTGGACTGTCGCCAGGCTG 791
OY 847 GTGTGATGTAGTGTGTTGGACTCTTCCAGGAAATCCATCTCCAGATCAATATGA 906
DB 792 GTGTGATGTAGTGTGTTGGACTCTTCCAGGAAATCCATCTTCCAGATCAATATGA 851
OY 907 TCAAGTACATCAAAAGCAATACCTAATCTCCAATCTATTGGAGGCAATGTGTCTAG 966
DB 852 TCAAGTACATCAAAAGCAATACCTAATCTCCAATCTATTGGAGGCAATGTGTCTAG 911
OY 967 CTGCCAGGCCAAGAACCTCATATTGATGACAGTGTGTGATCCCTGCGGTTGGCATGGAA 1026
DB 912 CTGCCAGGCCAAGAACCTCATATTGATGACAGTGTGTGATCCCTGCGGTTGGCATGGAA 971
```

OY	1027	GTCGCTCCATCTGCATTATCCAGAAAGTGCTGGCCTGTGGCGGCC	AAGCAACAGCAG	1086
Dd	972	GTGGCTCATCTGCATTTACGAAGAAGTGTGGCCCTGTGGCGGCC	AAGCAACAGCAG	1031
OY	1087	TGTACAAGGTATAGTAGTACGACGCGCCTTGTGTCTCCGGTCAAT	CTGATGAGGAA	1146
Dd	1032	TGTACAAGGTATGTAAGATATGACGCGCCTTGTGTCTCCGGTCAAT	CTGATGAGGAA	1091
OY	1147	TCCAATAATGGGTCATATTGGGAAGCCTTGCCCTTGCGGCCCTTC	CAGTCATGATGG	1206
Dd	1092	TCCAATAATGGGTCATATTGGGAAGCCTTGCCCTTGCGGCCCTTC	CAGTCATGATGG	1151
OY	1207	GCTCTCTCTGGCTSCAACACCACACTGAGGCCCTTGATTAATCTCT	CCGATGGGATCC	1266
Dd	1152	GCTCTCTCTGGCTSCCACCACTGAGGCCCTTGATTAATCTCTCT	CCGATGGGATCC	1211
OY	1267	GGCTTAAGAAAATATGCGCGTATGGTTCTCTGATGCGCATGACAAAG	ACCTCAGCAGCC	1326
Dd	1212	GGCTTAAGAAAATATGCGCGTATGGTTCTCTGATGCGCATGACAAAG	ACCTCAGCAGCC	1271
OY	1327	AGAACAGATATTTCACTAGTAAGCTGCACAAATCAAAGTGGCCAGGAA	TGTCTGSTGCTG	1386
Dd	1272	AGAACAGATATTTCACTAGTAAGCTGCACAAATCAAAGTGGCCAGGAA	TGTCTGSTGCTG	1331
OY	1387	TGCAGAGCAAAAGGTCATCCACAAATTTGTCCCTTACTGTATGTCT	GCATCCAAACT	1446
Dd	1332	TGCAGAGCAAAAGGTCATCCACAAATTTGTCCCTTACTGTATGTCT	GCATCCAAACT	1391
OY	1447	CATGCCAGACATTTGTGTGCCAAGACCTTGACCCAAAGTCCGAGCCATG	TGTACTGTGGGG	1506
Dd	1392	CATGCCAGACATTTGTGTGCCAAGACCTTGACCCAAAGTCCGAGCCATG	TGTACTGTGGGG	1451
OY	1507	AGCTTTAACTTTGAGAAAGAACGTCCTTCAGCCACAGGTGGAAAGSTGGC	TTCAT	1559
Dd	1452	AGCTTTAACTTTGAGAAAGAACGTCCTTCAGCCACAGGTGGAAAGSTGGC	TTCAT	1504
RESULT 4				
ID	ABV24746	standard; cdna: 1677 BP.		
XX	ABV24746:			
AC	ABV24746:			
XX	16-SEP-2002 (first entry)			
DT				
XX	Human prostate expression marker CDNA 24737.			
DE				
XX	Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker			
KW	pharmacogenomic marker; gene: ss.			
XX	Homo sapiens.			
OS	MO200160860-A2.			
PN				
PD	23-AUG-2001.			
XX				
Pf	20-FEB-2001: 2001WO-USO5171.			
XX				
PR	17-FEB-2000: 2000US-183319P.			
FR	16-MAR-2000: 2000US-189862P.			
PR	25-MAY-2000: 2000US-207454P.			
PR	08-JUN-2000: 2000US-211314P.			
PR	18-JUL-2000: 2000US-219007P.			
PR	13-DEC-2000: 2000US-255281P.			
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.			
PA	Schlegel R, Endege WO, Monahan JE;			
XI	WPI: 2001-662795/76.			
XX				
XR	Noel isolated nucleic acid molecule associated with			
XT	cancerous state of			

PT	prostatecells and correlating with presence of prostate cancer. useful
PT	for detecting presence of prostate cancer, stage of prostate cancer -
XX	
PS	Claim 1; Page 4726; 11750pp: English.
XX	
CC	The invention relates to an isolated nucleic acid molecule (1) comprising
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC	specification of its complement. (1) is useful for:
CC	(a) assessing whether a patient is afflicted with prostate cancer;
CC	(b) monitoring the progression of prostate cancer in a patient;
CC	(c) assessing the efficacy of a test compound to inhibit prostate
CC	cancer in a patient;
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC	in a patient;
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound;
CC	(g) determining whether prostate cancer has metastasized in a patient;
CC	(h) assessing the aggressiveness or indolence of prostate cancer in a
CC	patient;
CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX	
SQ	Sequence 1677 Bp: 428 A: 446 C: 445 G: 358 T: 0 other:
Query Match	55.3% Score 914: DB 23: Length 1677:
Best Local Similarity	99.3% Pred. No. 0:
Matches 1364: Conservative	Mismatches 9: Indels 0: Gaps 0:
OY	187 ACCAGGTGACCTGACTTTCGCTTGACCACAATAATCATCTTTAAGACCCCACTGGTTT 246
Db	132 ACCAGGTGACCTGACTTTCGCTTGACCACAATAATCATCTTTAAGACCCCACTGGTTT 191
OY	247 CCTTCCCATTGACACAGTCACAGAGCGTGGGATGGCCATACCATGGCGTTTACAGGC 306
Db	192 CCTTCTCCATTGACACAGTCACAGAGCGTGGGATGGCCATACCATGGCGTTTACAGGC 251
OY	307 GTATTGGCTTCATCCACCACAACCTGTACACCTGAATTTCCAGGCCAATGAAGTTCCGAAG 366
Db	252 GTATTGGCTTCATCCACCACAACCTGTACACCTGAATTTCCAGGCCAATGAAGTTCCGAAG 311
OY	367 TGAAGAATAATGAACAGGAGATTATCACAGACCCGTGGTCTCCAGGCCCAAGAGATCCGG 426
Db	312 TGAAGAATAATGAACAGGAGATTATCACAGACCCGTGGTCTCCAGGCCCAAGAGATCCGG 371
OY	427 TGCGGGATGTTTTAGAGCCCAAGGCCCGGCATGTTTTGCGGTAATCCCAATCAGACA 486
Db	372 TGCGGGATGTTTTAGAGCCCAAGGCCCGGCATGTTTTGCGGTAATCCCAATCAGACA 431
OY	487 CAGGCCCGATGGGAGCCGCTTGATGGAGCATATCTCCCTCCAGGGACATTTGATTTTCTCA 546
Db	432 CAGGCCCGATGGGAGCCGCTTGATGGAGCATATCTCCCTCCAGGGACATTTGATTTCTCA 491
OY	547 AAGAGAGAGAACATGACTGTTTCTTGAAAGAAATATATGACAAAGAGGGAAGCACTTGCG 606
Db	492 AAGAGAGAGAACATGACTGTTTCTTGAAAGAAATATATGACAAAGAGGGAAGCACTTGCG 551
OY	607 TAGCCCCCGCAGCATCACTGAAGAGGCAATGAATAATTGTGACGCCAGCAAGAAG 666
Db	552 TAGCCCCCGCAGCATCACTGAAGAGGCAATGAATAATTGTGACGCCAGCAAGAAG 611
OY	667 GAAGCTTCCCTCTTAAATGAAGATGATGACTTGTGGCCATCATTTGCCCGCCAGACACC 726
Db	612 GAAGCTTCCCTCTTAAATGAAGATGATGACTTGTGGCCATCATTTGCCCGCCAGACACC 671
OY	727 TGAAGAAGAAATCGGACATACCCACTAGCCTCCAAGATATGCCAAGAAACAGCTGCTGTGG 786
Db	672 TGAAGAAGAAATCGGACATACCCACTAGCCTCCAAGATATGCCAAGAAACAGCTGCTGTGG 731
OY	787 GGGCAGCATTGGCACTCATGAGATGACAAATATAGCTGTGACTTGCCTCGCCAGGCTG 846
Db	732 GGGCAGCATTGGCACTCATGAGATGACAAATATAGCTGTGACTTGCCTCGCCAGGCTG 791
OY	847 GTGTGATGTGTGTTTGGACTTCCCAAGGAAATTCATCTTCCAGATCATATATGA 906

Db 792 GTGTGATGATGATGTTTGGACTCTTCCAGAGGAAATTCATCTTCAGATCATATATGA 851  
QY 907 TCAAGTACATCAAGACAAATPACCTTATCTCCAGATCATTTGGAGCAATGGTCTCTG 966  
Db 852 TCACATACATCAAGACAAATPACCTTATCTCCAGATCATTTGGAGCAATGGTCTCTG 911  
QY 967 CTGCCCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGCTGCTGCGGGTGGAGGAA 1026  
Db 912 CTGCCCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGCTGCGGGTGGAGGAA 971  
QY 1027 GTGCTCTCATCTGATATATCCAGAAAGTGTGCTGTGGCGGCCCAAGCAACACACAG 1086  
Db 972 GTGCTCTCATCTGATATATCCAGAAAGTGTGCTGTGGCGGCCCAAGCAACACACAG 1031  
QY 1087 TGTACAAAGTGTATGATATGACAGCGGCTTTGGTGTTCGGCATTTGCGATGGAGGAA 1146  
Db 1032 TGTACAAAGTGTATGATATGACAGCGGCTTTGGTGTTCGGCATTTGCGATGGAGGAA 1091  
QY 1147 TCCAAAATGTGGGTCAATATTCGAAAGCCTTGGCCCTTCCACAGTCATGATG 1206  
Db 1092 TCCAAAATGTGGGTCAATATTCGAAAGCCTTGGCCCTTCCACAGTCATGATG 1151  
QY 1207 GCTCTCTCCCTGGTCCACACAGTGTGAGCCCTGTGATATCTTTCCGATGGAGTCC 1266  
Db 1152 GCTCTCTCCCTGGTCCACACAGTGTGAGCCCTGTGATATCTTTCCGATGGAGTCC 1211  
QY 1267 GGTAAAGAAATATCCGGGTATGGTCTCTCGATGCGATGACAGCACCCTCAGAGCC 1326  
Db 1212 GGTAAAGAAATATCCGGGTATGGTCTCTCGATGCGATGACAGCACCCTCAGAGCC 1271  
QY 1327 AGAAGACATATTTTCAGTGAAGTGCACAAATCAAAAGTGGCCAGAGAGTGTGCTGTG 1386  
Db 1272 AGAAGACATATTTTCAGTGAAGTGCACAAATCAAAAGTGGCCAGAGAGTGTGCTGTG 1331  
QY 1387 TCGAGAGCAAGGGTCATATCCAAATTTTCCCTTACCTGATTTGGGATCCAAACT 1446  
Db 1332 TCGAGAGCAAGGGTCATATCCAAATTTTCCCTTACCTGATTTGGGATCCAAACT 1391  
QY 1447 CATGCCAGACATTTGGTCCAGAGCTTGAAGCCAGTCCAGGCATGATGATCTGTGGG 1506  
Db 1392 CATGCCAGACATTTGGTCCAGAGCTTGAAGCCAGTCCAGGCATGATGATCTGTGGG 1451  
QY 1507 AGCTTAAGTTTGAAGAGAGAACGTCCTCAGGCCAGGTGGAAGGTGGCTCCAT 1559  
Db 1452 AGCTTAAGTTTGAAGAGAGAACGTCCTCAGGCCAGGTGGAAGGTGGCTCCAT 1504

RESULT 5  
ABK34910  
ID ABK34910 standard; cDNA; 1157 BP.  
XX  
AC ABK34910;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Human cDNA encoding secreted protein #48.  
XX  
XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KM viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KM autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KM autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KM Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;  
KM coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
KM tissue regeneration; wound healing; burn; haematopoiesis;  
KM myeloid cell deficiency; lymphoid cell deficiency.

OS Homo sapiens.  
XX  
PN WO200177288-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 29-MAR-2001; 2001WO-US10224.

XX  
PR 06-APR-2000; 2000US-195582P.  
XX  
PA (GEMT ) GENETICS INST INC.  
XX  
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;  
PI Gulukota K, Graham JR;  
XX  
DR WPI; 2002-179321/23.  
XX  
PT Five hundred and ninety two polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT

Claim 1; Page 93; 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins. The polynucleotides can be used as probes for the  
CC identification and isolation of full length cDNA and genomic DNA. The  
CC polynucleotides and proteins can also be used as nutritional supplements.  
CC The proteins are useful in the treatment of various immune deficiencies  
CC and disorders such as viral infections, bacterial infections, fungal  
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
CC and conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment of  
CC burns, incisions and ulcers. The proteins are also useful for regulating  
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention.  
XX

Sequence 1157 BP; 297 A; 270 C; 326 G; 264 T; 0 other;

Query Match 49.7%; Score 822; DB 24; Length 1157;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1022; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 619 GCATCAGCTGAAGAGGAGCAATGAATTTCTGCAGCGCAGACAGAGGAAAGTGGCCA 678  
Db 131 GCATCAGCTGAAGAGGAGCAATGAATTTCTGCAGCGCAGACAGAGGAAAGTGGCCA 190  
QY 679 TTGTAATGAAGATGATGACCTTGTGGCATATTTGCCGAGACAGACTTAAGAAGATC 728  
Db 191 TTGTAATGAAGATGATGACCTTGTGGCATATTTGCCGAGACAGACTTAAGAAGATC 250  
QY 739 GGGACTACCCAGCTAGCCTCCAAAGATGCCAAGAAACAGCTGTGTGGGGCAGCCATTG 798  
Db 251 GGGACTACCCAGCTAGCCTCCAAAGATGCCAAGAAACAGCTGTGTGGGGCAGCCATTG 310  
QY 799 GCACCTATGAGATGACAAATATAGGTGTGACCTTGTGCCCGAGGCTGTGTGATGTAG 858  
Db 311 GCACCTATGAGATGACAAATATAGGTGTGACCTTGTGCCCGAGGCTGTGTGATGTAG 370  
QY 859 TGGTTTGGACCTCTCCAGAGGAAATTCATCTCCAGATCATATGATCAATGATCA 918  
Db 371 TGGTTTGGACCTCTCCAGAGGAAATTCATCTCCAGATCATATGATCAATGATCA 430  
QY 919 AAGACAAATACCTTAATCTTCAAGTCAATTTGAGAGCATGTGTCTGCTGCGCCAGGCCA 978  
Db 431 AAGACAAATACCTTAATCTTCAAGTCAATTTGAGAGCATGTGTCTGCTGCGCCAGGCCA 490  
QY 979 AGAAGCTCATTTGATGACAGTGTGATGCCCTGCGGGTGGGATGGAGTGGCTCCATCT 1038  
Db 491 AGAAGCTCATTTGATGACAGTGTGATGCCCTGCGGGTGGGATGGAGTGGCTCCATCT 550  
QY 1039 GCATTTACGAGAGAGTGTGGCTGTGGGGGCCCAAGCAAGCAGTGTACAAAGTGT 1098  
Db 551 GCATTTACGAGAGAGTGTGGCTGTGGGGGCCCAAGCAAGCAGTGTACAAAGTGT 610



QY	1439	CCAACACTATGCGAGACATTGGTCCCAAGACACTTGACCCAAAGTCGAGCCATGATGTA	1498
Db	1675	CCAAACACTATGCGAGACATTGGTCCCAAGACACTTGACCCAAAGTCGAGCCATGATGTA	1734
QY	1499	CTCTGGGAGGCTTAAGTTTGAGAGAAGACGTCCTCAGCCAGGTGGAGTGGCCGCCA	1558
Db	1735	CTCTGGGAGGCTTAAGTTTGAGAGAAGACGTCCTCAGCCAGGTGGAGTGGCCGCCA	1794
QY	1559	TAGCCTCATTGCGATGAGAACGGCGCTTTCTGAAAAAGGATCCAGCACACTCTCTGGT	1618
Db	1795	TAGCCTCATTGCGATGAGAACGGCGCTTTCTGAAAAAGGATCCAGCACACTCTCTGGT	1854
QY	1619	TTTTTTTTCATATAAAGTTTAGAACA	1645
Db	1855	TTTTTTTTCATATAAAGTTTAGAACA	1881
RESULT 7			
ID	AAF21634	AAF21634 standard; DNA; 2019 BP.	
XX	XX	AAF21634;	
AC	XX	27-MAR-2001 (first entry)	
DT	XX	Human breast and ovarian cancer associated antigen gene SEQ ID 21.	
DE	XX	Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;	
KM	XX	neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
KM	XX	antidiabetic; antinflammatory; antitumor; antiparasitic; anticonvulsant;	
KM	XX	antibacterial; antifungal; antiparasitic; cardiant; immune disorder;	
KM	XX	Addison's disease; allergy; autoimmune haemolytic anaemia;	
KM	XX	autoimmune thyroiditis; diabetes mellitus; Crohn's disease;	
KM	XX	multiple sclerosis; rheumatoid arthritis; ulcerative colitis;	
KM	XX	cardiovascular disorder; wound healing; neurological disease; ds.	
OS	XX	Homo sapiens.	
PN	XX	WO200055173-A1.	
PD	XX	21-SEP-2000.	
PF	XX	08-MAR-2000; 2000WO-US05881.	
PR	XX	12-MAR-1999; 99US-0124270.	
XX	XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	XX	Rosen CA, Ruben SM;	
PI	XX	WPI; 2000-611515/58.	
DR	XX	P-PSDB; AAB58731.	
PT	XX	New human breast and ovarian cancer associated gene sequences and the	
PT	XX	polypeptides encoded by these genes, useful in the prevention,	
PT	XX	treatment and diagnosis of cancer, immune disorders, cardiovascular	
PT	XX	disorders and neurological diseases -	
PS	XX	Claim 1: Page 496: 1299pp; English.	
CC	XX	Sequences AAF21614 - AAF22031 represent DNA sequences encoding human	
CC	XX	proteins AAB58711 - AAB59128. The DNA and protein sequences are	
CC	XX	associated with breast and ovarian cancer. Included in the invention are	
CC	XX	sequences AAF22032 - AAF22040 and AAB59129 which are used in the	
CC	XX	isolation and characterisation of the DNA and protein sequences of the	
CC	XX	invention. The breast and ovarian cancer associated DNA, protein, agonist	
CC	XX	or antagonist sequences exhibit cytostatic; immunosuppressive;	
CC	XX	neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
CC	XX	antibacterial; antinflammatory; antitumor; antiparasitic; anticonvulsant;	
CC	XX	antidiabetic; antifungal; antiparasitic and cardiant activity. The	
CC	XX	polynucleotide and protein sequences are used in the diagnosis of cancer,	
CC	XX	particularly breast and ovarian cancer. The nucleic acid sequences,	
CC	XX	proteins, agonists and agonists may also be used in the diagnosis,	

CC	prevention and treatment of immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular disorders such as myocardial ischaemia; wound healing; neurological diseases such as cerebral anoxia and epilepsy; and infectious diseases.
XX	Sequence 2019 BP; 512 A; 483 C; 538 G; 476 T; 10 other:
Query Match	40.0%; Score 662; DB 21; Length 2019;
Best Local Similarity	99.5%; Pred. No. 6,8e-306;
Matches 982;	Conservative 0; Mismatches 4; Indels 1; Gaps 1
QY	659 CAGAAGGAAAGTTGCCCATGTTAATGAGATGATGAGCTTGGCCATATTGCCCG 718
DB	985 CAAGAAGGAAAGTTGCCCATGTTAATGAGATGATGAGCTTGGCCATATTGCCCG 1044
QY	719 GACAGACTGAAGAAGATCGGGACTACCCACTAGCCTCCAAAGATGCCAAGAACGT 778
DB	1045 GACAGACCTGGAAGAAGAAATCGGGACTACCCACTAGCCTCCAAAGATGCCAAGAACGT 1104
QY	779 GGTGTGTGGGGAGAGCATTTGGGACATGAGGATGACAGTATPAGGCTGAGCTGGCC 838
DB	1105 GCTGTGTGGGGAGAGCATTTGGGACATGAGGATGACAGTATPAGGCTGAGCTGGCC 1164
QY	839 CCAGCGTGTGTGATGTAGTGGTTTTGGACCTTCCAGGGGAAATTCATCTTCAGAT 898
DB	1165 CCAGCGTGTGTGATGTAGTGGTTTTGGACCTTCCAGGGGAAATTCATCTTCAGAT 1224
QY	899 CAATATGATCAAGTACATCAAAACAAATACCTTAATCTCCAAAGTCATTTGAGCAGATCT 958
DB	1225 CAATATGATCAAGTACATCAAAACAAATACCTTAATCTCCAAAGTCATTTGAGCAGATCT 1284
QY	959 GGTCACTGCTGGCCAGGCCAAGACCTCAATTCAGTACAGTGTGAGATGACCTCGGGTGG 1018
DB	1285 GGTCACTGCTGGCCAGGCCAAGACCTCAATTCAGTACAGTGTGAGATGACCTCGGGTGG 1344
QY	1019 CATGGGAATGGCTCCATCTGCATTTATCCAGAGAGTGTGGCTGTGGGCGGCCCAAC 1078
DB	1345 CATGGGAATGGCTCCATCTGCATTTATCCAGAGAGTGTGGCTGTGGGCGGCCCAAC 1404
QY	1079 AACAGCAGTGTCAAAAGTGTATGAGTATGCCAGCGGCTTTGGTGTTCGGGTCATTGTCTGA 1138
DB	1405 AACAGCAGTGTCAAAAGTGTATGAGTATGCCAGCGGCTTTGGTGTTCGGGTCATTGTCTGA 1464
QY	1139 TGGAGGAATCCAAATATGGGTCATATTGGCGAAAGCTTGGCCCTTGGGCGCTCCACACT 1198
DB	1465 TGGAGGAATCCAAATATGGGTCATATTGGCGAAAGCTTGGCCCTTGGGCGCTCCACACT 1523
QY	1199 CATGATGGGCTCTCTCTCGGCTGGCCACACTGATAGGCCCTTGGTGAATATCTTTTTCGA 1258
DB	1524 CATGATGGGCTCTCTCTCGGCTGGCCACACTGATAGGCCCTTGGTGAATATCTTTTTCGA 1583
QY	1259 TGGGATCCGGCTAAAGAAATATCGCGATATGGGTTCTCTCGATGCCATGAGCAAGCACT 1318
DB	1584 TGGGATCCGGCTAAAGAAATATCGCGATATGGGTTCTCTCGATGCCATGAGCAAGCACT 1643
QY	1319 CAGCAGCCGAAGACAGATATTTTCAATGAGAGCTGACAAATCAAGATGGCCGACGAGAGTGC 1378
DB	1644 CAGCAGCCGAAGACAGATATTTTCAATGAGAGCTGACAAATCAAGATGGCCGACGAGAGTGC 1703
QY	1379 TGGTCTCTGCGAGAGCAAAAGGCTCAATCCAAATTTTGGCTTACCTGATTTGGCGCAT 1438
DB	1704 TGGTCTCTGCGAGAGCAAAAGGCTCAATCCAAATTTTGGCTTACCTGATTTGGCGCAT 1763
QY	1439 CCAGACCTCATCCAGAGACATTTGGTCCAAAGAGCTTGAGCCCAAGTCCGAGCCATGATGA 1498
DB	1764 CCAGACCTCATCCAGAGACATTTGGTCCAAAGAGCTTGAGCCCAAGTCCGAGCCATGATGA 1823
QY	1499 CTCCTGGGAGCTTAAAGTTTGAGAAGAGAACTTCCTCAGCCCAAGGTGGAAGGTGGCTCCA 1558
DB	1824 CTCCTGGGAGCTTAAAGTTTGAGAAGAGAACTTCCTCAGCCCAAGGTGGAAGGTGGCTCCA 1883





FT /\*tag= a  
TT /product= "Human type II IMPDH-AGRP variant"  
XX  
XX WO200185952-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 10-MAY-2001; 2001WO-US15457.  
XX  
XX 10-MAY-2000; 2000US-203448P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Krysiek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff Jr, Einspahr HM, Kish K;  
PI  
XX WPI: 2002-164105/21.  
XX P-PSDB: AAE18181.  
XX  
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies  
XX  
XX Claim 54; Fig 14; 161pp: English.  
XX  
XX The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH: EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding antibodies,  
CC as diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human type II IMPDH-AGRP variant cDNA.  
XX  
XX Sequence 1158 BP: 277 A; 300 C; 313 G; 268 T; 0 other;  
SQ  
Query Match 37.1%; Score 613; DB 24; Length 1158;  
Best Local Similarity 99.5%; Pred. No. 1,7e-282;  
Matches 813; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
XX  
XX 776 GCTGCTGTGGGCGAGCCATTGGCAGCTCATGAGAGTACAAAGTATAGCTGACTTGCT 835  
XX  
XX 342 GCTGCTGTGTGGGCGAGCCATTGGCAGCTCATGAGAGTACAAAGTATAGCTGACTTGCT 401  
XX  
XX 836 CGCCAGGCTGGTGGAGTGTAGTGGTTTGGAGCTTCCAGGAAATTCATCTTCCA 895  
XX  
XX 402 CGCCAGGCTGGTGGAGTGTAGTGGTTTGGAGCTTCCAGGAAATTCATCTTCCA 461  
XX  
XX 896 GATCAATATGATCAAGTACATCAAGCAAAATACCTATATCCAAAGTCATTGGAGCAA 955  
XX  
XX 462 GATCAATATGATCAAGTACATCAAGCAAAATACCTATATCCAAAGTCATTGGAGCAA 521  
XX  
XX 956 TGTGCTCACTGCTGCCAGGCGCAAGAACCTGATTGATGACAGTGTGATGCCCTGGGGGT 1015  
XX  
XX 522 TGTGCTCACTGCTGCCAGGCGCAAGAACCTGATTGATGACAGTGTGATGCCCTGGGGGT 581  
XX  
XX 1016 GGGGATGGAGTGGCTCCATCTGCATATATCCAGGAATGCTGCCCTTGGGCGGCCCA 1075  
XX  
XX 582 GGGGATGGAGTGGCTCCATCTGCATATATCCAGGAATGCTGCCCTTGGGCGGCCCA 641  
XX  
XX 1076 AGCAACAGCACTGTACAAAGTGTATGATATGACAGGCGCTTGTGTTCCGGTCATTGC 1135  
XX  
XX 642 AGCAACAGCACTGTACAAAGTGTATGATATGACAGGCGCTTGTGTTCCGGTCATTGC 701  
XX  
XX 1136 TGATGAGAGGAATCCAAATGTGGGTCAATATTCGAAAGCCTTGGCCCTTGGGCGCTCAC 1195

Db |||||  
Db 702 TGATGAGAGGAATCCAAATGTGGGTCAATATTCGAAAGCCTTGGCCCTTGGGCGCTCAC 761  
XX  
XX 1196 AGTCATGATGGGCTCTCCGCTGCGTCCACACATGAGAGCCCTGGTAATACCTTTTC 1255  
XX  
XX 762 AGTCATGATGGGCTCTCTCCGCTGCGTCCACACATGAGAGCCCTGGTAATACCTTTTC 821  
XX  
XX 1256 CGATGAGATCCGGCTAAAGAAATATCCGCGTATGCGTCTCTCATGATGCAAGCA 1315  
XX  
XX 822 CGATGAGATCCGGCTAAAGAAATATCCGCGTATGCGTCTCTCATGATGCAAGCA 881  
XX  
XX 1316 CCTCAGAGCCAGCAAGATATTTCACTGAAGCTGACAAATCAATCAAGTGGCCAGSAGT 1375  
XX  
XX 882 CCTCAGAGCCAGCAAGATATTTCACTGAAGCTGACAAATCAATCAAGTGGCCAGSAGT 941  
XX  
XX 1376 GTCGTGCTGCTGACAGCAAGGCTCAATCCAAATTTTCCCTTACCTGATTTGCTGG 1435  
XX  
XX 942 GTCGTGCTGCTGACAGCAAGGCTCAATCCAAATTTTCCCTTACCTGATTTGCTGG 1001  
XX  
XX 1436 CATCAGACATCATGCCAGACATTTGCTGCCAAGAGCTTGACCAATCCGAGCATGAT 1495  
XX  
XX 1002 CATCAGACATCATGCCAGACATTTGCTGCCAAGAGCTTGACCAATCCGAGCATGAT 1061  
XX  
XX 1496 GTACTGTGGGAGCTTAAGTTTGAGAAAGAACGCTCTCAGGCCAGGTGGAAGGTGGCGT 1555  
XX  
XX 1062 GTACTGTGGGAGCTTAAGTTTGAGAAAGAACGCTCTCAGGCCAGGTGGAAGGTGGCGT 1121  
XX  
XX 1556 CCATAGGCTTCATTCGTATGAGAAAGCGGCTTTCTGA 1592  
XX  
XX 1122 CCATAGGCTTCATTCGTATGAGAAAGCGGCTTTCTGA 1158  
XX  
XX RESULT 10  
XX AAD28922 standard; cDNA; 1155 BP.  
XX  
XX AAD28922;  
XX  
XX 07-MAY-2002 (first entry)  
XX  
XX Human type II IMPDH-DKT variant cDNA.  
XX  
XX Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;  
XX KW proliferative-type disease; cancer; imaging methodology; cytostatic;  
XX KM therapy; enzyme; variant; ss.  
XX  
XX Homo sapiens.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX FH 1..1155  
XX FT CDS  
XX FT /tag= a  
XX FT /product= "Human type II IMPDH-DKT variant"  
XX  
XX WO200185952-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 10-MAY-2001; 2001WO-US15457.  
XX  
XX PF 10-MAY-2000; 2000US-203448P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Krysiek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
XX PI Mouravieff Jr, Einspahr HM, Kish K;  
XX  
XX WPI: 2002-164105/21.  
XX  
XX P-PSDB: AAE18181.  
XX  
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating

PT antibodies -  
XX  
PS Claim 54: Fig 5; 161pp; English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human type II IMPDH-DKT variant cDNA.  
XX  
SQ Sequence 1155 BP; 281 A; 299 C; 310 G; 265 T; 0 other;  
XX  
Query Match 37.0%; Score 612; DB 24; Length 1155;  
Best Local Similarity 99.5%; Pred. No. 5.2e-282;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
XX  
QY 777 CTCTGTGTGGGCGACCCATTGGCCTCATGAGATGACAGATATAGCTTGACCTTGCTC 836  
DB 340 CTCTGTGTGGGCGACCCATTGGCCTCATGAGATGACAGATATAGCTTGACCTTGCTC 399  
QY 837 GCCCAGGCTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 896  
DB 400 GCCCAGGCTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 459  
QY 897 ATCAATATGATCAAGTATGATCAAGTATGATCAAGTATGATCAAGTATGATCAAGTATGAT 956  
DB 460 ATCAATATGATCAAGTATGATCAAGTATGATCAAGTATGATCAAGTATGATCAAGTATGAT 519  
QY 957 GTGGTACCTGCTGCCAGGCGAAGCCTCATGATGATGATGATGATGATGATGATGATGATGAT 1016  
DB 520 GTGGTACCTGCTGCCAGGCGAAGCCTCATGATGATGATGATGATGATGATGATGATGATGAT 579  
QY 1017 GGCATGGGAAGTGGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATGCTCAT 1076  
DB 580 GGCATGGGAAGTGGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATGCTCATGCTCAT 639  
QY 1077 GCAACAGCGGTGTACAAGGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1136  
DB 640 GCAACAGCGGTGTACAAGGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 699  
QY 1137 GATGAGGAATCCAAATGTGGTCAATGTGGCAAGCCTTGCGGCTCCACCA 1196  
DB 700 GATGAGGAATCCAAATGTGGTCAATGTGGCAAGCCTTGCGGCTCCACCA 759  
QY 1197 GTATATATGGGCTCTCTCTGGCTGCCACGACGAGGCCCTTGATTAATCTTTTTC 1256  
DB 760 GTATATATGGGCTCTCTCTGGCTGCCACGACGAGGCCCTTGATTAATCTTTTTC 819  
QY 1257 GATGAGATCCGCTTAAGAATATCCCGGATGATGATGATGATGATGATGATGATGATGATGAT 1316  
DB 820 GATGAGATCCGCTTAAGAATATCCCGGATGATGATGATGATGATGATGATGATGATGATGAT 879  
QY 1317 CTGAGCAGCCAGACAGATATTTTCACTGAAGCTGACCAAAATCAAACTGACCCAGGAGTG 1376  
DB 880 CTGAGCAGCCAGACAGATATTTTCACTGAAGCTGACCAAAATCAAACTGACCCAGGAGTG 939  
QY 1377 TCTGTGTGTGTGAGAGCAAAAGGTCATCATCAAAATTTTCTCCCTTACCTGATGCTGGC 1436  
DB 940 TCTGTGTGTGTGAGAGCAAAAGGTCATCATCAAAATTTTCTCCCTTACCTGATGCTGGC 999  
QY 1437 ATCAACACTCATGACGAGCATTTGCTGCCAGAGCTTGACCCCAAGCTGCAGCATGATG 1496  
DB 1000 ATCAACACTCATGACGAGCATTTGCTGCCAGAGCTTGACCCCAAGCTGCAGCATGATG 1059

QY 1497 TACTCTGGGAGCTTAAGTTTGAGAGAGACGTCCTCAGCCAGGTGGAAGTGGCGTC 1556  
DB 1060 TACTCTGGGAGCTTAAGTTTGAGAGAGAACGTCCTCAGCCAGGTGGAAGTGGCGTC 1119  
QY 1557 CATAGCTCCATTCGATGAGAGCGGCTTTTCTGA 1592  
DB 1120 CATAGCTCCATTCGATGAGAGCGGCTTTTCTGA 1155  
RESULT 11  
AAD28923  
ID AAD28923 standard; cDNA: 1155 BP.  
AC AAD28923;  
DT 07-MAY-2002 (first entry)  
XX Human type II IMPDH-SPS variant cDNA.  
DE  
XX Human; inosine 5'-monophosphate dehydrogenase: IMPDH: drug discovery;  
KM proliferative-type disease: cancer; imaging methodology; cytostatic;  
KW therapy; enzyme; variant; ss.  
OS Homo sapiens.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT CDS 1..1155  
FT /\*tag= a  
FT /product= "Human type II IMPDH-SPS variant"  
PN MO200185952-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 10-MAY-2001; 2001MO-US15457.  
XX  
PR 10-MAY-2000; 2000US-203448P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Kryslek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Moutavieff JE, Einspahr HM, Kish K;  
XX  
DR MPI: 2002-164105/21.  
DR P-PSDB: AAEL8167.  
XX  
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
PS Claim 54: Fig 7; 161pp; English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human type II IMPDH-SPS variant cDNA.  
XX  
SQ Sequence 1155 BP; 278 A; 300 C; 310 G; 267 T; 0 other;

Query Match	37.0%	Score 612	DB 24	Length 1155	
Best Local Similarity	99.5%	Pred. No. 5.2e-282			
Matches 812: Conservative	0	Mismatches 4	Indels	0	Gaps
OY	777	CTGCTGTGTGGGGCGACGACCTTGGACCTCATGATGAGATGACAACTATATAGGCTGGACTTGGCTC	836		
DB	340	CTGCTGTGTGGGGCGACGACCTTGGACCTCATGAGATGACAACTATATAGGCTGGACTTGGCTC	399		
OY	837	GCCACGGCTGTGTGGATGTAGTGGTTTTGGACCTCTTCCACAGAAATTCATCTTCCAG	896		
DB	400	GCCACGGCTGTGTGGATGTAGTGGTTTTGGACCTCTTCCACAGAAATTCATCTTCCAG	459		
OY	897	ATCAATATGATCAATAGTACATCTCAAGCAAAATACCTATATCTCAAGTCAATTGGAGCAAT	956		
DB	460	ATCAATATGATCAATAGTACATCTCAAGCAAAATACCTATATCTCAAGTCAATTGGAGCAAT	519		
OY	957	GTGGTACATGCTGGCCCAAGGCAAGCAACCATGTGATGCGAGGTGTGGATGCCCCGTGGGGGTG	1016		
DB	520	GTGGTACATGCTGGCCCAAGGCAAGCAACCATGTGATGCGAGGTGTGGATGCCCCGTGGGGGTG	579		
OY	1017	GCCATGTGGAAATGGGCTCCATCTGCATTTATCCAGAAATGCTGGCCCTGTGGGGCGGCCCAA	1076		
DB	580	GCCATGTGGAAATGGGCTCCATCTGCATTTATCCAGAAATGCTGGCCCTGTGGGGCGGCCCAA	639		
OY	1077	GCAACACCACTGTACAAAGTGTATGAGTATGACAGGCGCTTTGGTTCGGTCAATTGCT	1136		
DB	640	GCAACACCACTGTACAAAGTGTATGAGTATGACAGGCGCTTTGGTTCGGTCAATTGCT	699		
OY	1137	GATGAGGAGATCCAAAAATGTGGTCAATATTTGCGAAAGCCTTGCGCCCTTGGGGCGTCCACA	1196		
DB	700	GATGAGGAGATCCAAAAATGTGGTCAATATTTGCGAAAGCCTTGCGCCCTTGGGGCGTCCACA	759		
OY	1197	GTCATGATGGGCTCTCTCTGGCTGCCACACCTGAGGCGCCCTGGTGAATACTCTTTTTC	1256		
DB	760	GTCATGATGGGCTCTCTCTGGCTGCCACACCTGAGGCGCCCTGGTGAATACTCTTTTTC	819		
OY	1257	GATGGGATCCGGCTTAAGAAATATGCGGTATGGTCTCTCGATGCCATGGACAAAGCAC	1316		
DB	820	GATGGGATCCGGCTTAAGAAATATGCGGTATGGTCTCTCGATGCCATGGACAAAGCAC	879		
OY	1317	CTCAGCAGCGCAACAGATATTTCAGTGAAGCTGACCAAAATCAAAAGTGGCCAGGAGATG	1376		
DB	880	CTCAGCAGCGCAACAGATATTTCAGTGAAGCTGACCAAAATCAAAAGTGGCCAGGAGATG	939		
OY	1377	TCTGTGCTGTGGCAGACAAAGGATCAATCCCAAAATTTGTCCCTTACCTGATTTCTGTGC	1436		
DB	940	TCTGTGCTGTGGCAGACAAAGGATCAATCCCAAAATTTGTCCCTTACCTGATTTCTGTGC	999		
OY	1437	ATCCAACTCATATGCCAGACATTGGTGCACAGAGCTTGACCCAGTCCGAGCCATGATG	1496		
DB	1000	ATCCAACTCATATGCCAGACATTGGTGCACAGAGCTTGACCCAGTCCGAGCCATGATG	1059		
OY	1497	TACTGTGGGACCTTAAGTTTGAAGAAGAACGTCTCAGCCGACAGTGAAGCTGGCGTC	1556		
DB	1060	TACTGTGGGACCTTAAGTTTGAAGAAGAACGTCTCAGCCGACAGTGAAGCTGGCGTC	1119		
OY	1557	CATPACCTTCATTCGTATGAGAAGGCGCTTTCTGA	1592		
DB	1120	CATPACCTTCATTCGTATGAGAAGGCGCTTTCTGA	1155		
RESULT 12					
AAD28924					
ID	AAD28924 standard; cDNA; 1155 BP.				
XX	AAD28924;				
AC	AAD28924;				
XX	07-MAY-2002 (first entry)				
DT	Human type II IMPDH-SPT variant cDNA.				
XX	Human type II IMPDH-SPT variant cDNA.				
DE	Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;				
KW	Proliferative-type disease; cancer; imaging methodology; cytostatic;				

KM		Therapy; enzyme; variant; ss.
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1155
FT		/tag= a
FT		/product= "Human type II IMPDH-SPT variant"
XX		
PN	MO200185952-A2.	
XX		
PD	15-NOV-2001.	
XX		
PF	10-MAY-2001; 2001WO-US15457.	
XX		
PR	10-MAY-2000; 2000US-203448P.	
XX		
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.	
XX		
PI	Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;	
PI	MouraviJeff JE, Einspahr HM, Kish K;	
XX		
DR	WPI: 2002-164105/21.	
XX	P-PSDB: AAEL18172.	
PT	New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide	
PT	having an oligo-peptide domain substituted for a subdomain of a	
PT	wild-type IMPDH polypeptide, useful in drug discovery or for generating	
PT	antibodies -	
XX		
PS	Claim 54; Fig 11; 161pp; English.	
XX		
CC	The invention relates to modified inosine 5'-monophosphate dehydrogenase	
CC	(IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain	
CC	substituted for a subdomain of a wild-type IMPDH polypeptide. The	
CC	modified IMPDH polypeptides are useful for drug discovery, for	
CC	therapeutic, diagnostic and prognostic procedures for detecting or	
CC	quantifying modified IMPDH polypeptides and their corresponding nucleic	
CC	acids. IMPDH polypeptides are also useful for generating antibodies, as	
CC	diagnostic and prognostic markers of diseases, as targets for various	
CC	therapeutic modalities, and to identify and isolate ligands and other	
CC	agents that bind to modified IMPDH. These antibodies may be used in	
CC	diagnostic assays, imaging methodologies, therapeutic methods in the	
CC	management of cancer or other proliferative-type diseases, and in	
CC	purifying modified IMPDH polypeptides and for isolating related	
CC	molecules such as wild type and mutant IMPDH polypeptides. The present	
CC	sequence is human type II IMPDH-SPT variant cDNA.	
XX		
SO	Sequence 1155 BP; 278 A; 300 C; 309 G; 268 T; 0 other:	
	Query Match	37.0%; Score 612; DB 24; Length 1155;
	Best Local Similarity	99.5%; Pred. No. 5.2e-282;
	Matches 812; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
OY	777 CTGCTGTGTGGGGCAGCCATTGGCACTCATAGATGACAAAGTATAGCGTAGACTTGCTC	836
Dd	340 CTGCTGTGTGGGGCAGCCATTGGCACTCATAGATGACAAAGTATAGCGTAGACTTGCTC	399
OY	837 GCCAAGGCTGTGTGATGTAGTGGTTTGGACITCTTCCCAGGGAATTCCATCTTCAG	896
Dd	400 GCCCAGGCTGTGTGATGTAGTGGTTTGGACITCTTCCCAGGGAATAATTCATCTTCAG	459
OY	897 ATCATATATGATCAATACATCAAGACAANAACCCTAATCTCCAAGTCATTTGGAGGCAT	956
Dd	460 ATCATATATGATCAATACATCAAGACAANAACCCTAATCTCCAAGTCATTTGGAGGCAT	519
OY	957 GTGGTCACCTGTGCCCGAGCCAGAAGCTCAATGATGACGTGGATGGCCCTCGGGTG	1016
Dd	520 GTGGTCACCTGTGCCCGAGCCAGAAGCTCAATGATGACGTGGATGGCCCTCGGGTG	579
OY	1017 GGCAATGGGAAGTGGCTCCATCTGCATTATCCAGGAAGTGTGGCCTGTGGCGGCCCAA	1076

Db	580	GGCATGGCAAGTGGCTCCATCTGTCATTACGCAGAGAAAGTCTTG6GCTG	GGCGGCGCCCA	639
Qy	1077	GCACACAGCAGTGTACAGAGTGTATGAGTATGCAAGCGCGCTTGGTGT	CCGCGCATTTGCT	1136
Db	640	GCACACAGCAGTGTACAGAGTGTACAGAGTGTATGCAAGCGCGCTTGGTGT	CCGCGCATTTGCT	699
Qy	1137	GATGAGCAATCCAAATGTGGGTCATATTGGCAAAAGCCTTGGGCT	GGGCGCTCCACA	1196
Db	700	GATGAGCAATCCAAATGTGGGTCATATTGGCAAAAGCCTTGGGCT	GGGCGCTCCACA	759
Qy	1197	GTCAATGATGGGCTCTCTCTGGCTGCCACACAGTAGAGCCCGTGATGA	TACTTCTTTTCC	1256
Db	760	GTCAATGATGGGCTCTCTCTGGCTGCCACACAGTAGAGCCCGTGATGA	TACTTCTTTTCC	819
Qy	1257	GATGGGATCCGGCTTAAAGAAATATCCGGGTATGGGTTCTCTGATAGC	ATGAGACAAGAC	1316
Db	820	GATGGGATCCGGCTTAAAGAAATATCCGGGTATGGGTTCTCTGATAGC	ATGAGACAAGAC	879
Qy	1317	CTCAGCAGCCAGAAACAGATATTTCAGTGAAGCTGACCAAAATCAAGT	GCCAGAGAGTG	1376
Db	880	CTCAGCAGCCAGAAACAGATATTTCAGTGAAGCTGACCAAAATCAAGT	GCCAGAGAGTG	939
Qy	1377	TCTGTGCTGTGGCGAGCAAAAGGCTCAATCCCAAAATTTGTCCTTA	CTGATTTGCTGGC	1456
Db	940	TCTGTGCTGTGGCGAGCAAAAGGCTCAATCCCAAAATTTGTCCTTA	CTGATTTGCTGGC	999
Qy	1437	ATCCACACACTCATGCCAGAGACATTTGTCGCCAAGCTTGACCCCAAGT	CGAGCCATGATG	1496
Db	1000	ATCCACACACTCATGCCAGAGACATTTGTCGCCAAGCTTGACCCCAAGT	CGAGCCATGATG	1059
Qy	1497	TACTCTGGGAGCTTAAGTTTGAGAGAGAGACGCTCTCGACCCAGAGT	GAAGTTGGGCTG	1556
Db	1060	TACTCTGGGAGCTTAAGTTTGAGAGAGAGACGCTCTCGACCCAGAGT	GAAGTTGGGCTG	1119
Qy	1557	CATAGCCTCCATTGCTATGAGAAAGCGGCTTTTCTGA	1592	
Db	1120	CATAGCCTCCATTGCTATGAGAAAGCGGCTTTTCTGA	1155	
RESULT 13				
ID	AAD28925	standard: cDNA: 1155 BP.		
AC	AAD28925:			
XX	07-MAY-2002	(first entry)		
DT	07-MAY-2002	(first entry)		
XX	Human type II IMPDH-GSG variant cDNA.			
DE	Human type II IMPDH-GSG variant cDNA.			
XX	Human: inosine 5'-monophosphate dehydrogenase: IMPDH	drug discovery:		
KW	proliferative-type disease; cancer; imaging methodolgy; cytosolic;			
KW	therapy; enzyme; variant; ss.			
XX	Homo sapiens.			
OS	synthetic.			
XX	Key	Location/Qualifiers		
FT	CDS	1..1155		
FT		/tag- a		
FT		/product= "Human type II IMPDH-GSG variant"		
PN	W0200185952-A2.			
XX	15-NOV-2001.			
XX	10-MAY-2001: 2001WO-US15457.			
XX	10-MAY-2000: 2000US-203448P.			
XX	(BRIM ) BRISTOL-MYERS SQUIBB CO.			
XX	Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL,	Yan N:		
XX	Mouravieff JE, Einspahr HM, Kish K:			

XX	WPI: 2002-164105/21.
DR	P-PSDB: AAEI8174.
XX	
PT	New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
PT	having an oligo-peptide domain substituted for a subdomain of a
PT	wild-type IMPDH polypeptide, useful in drug discovery or for generating
PT	antibodies -
XX	
PS	Claim 54; Fig 9; 16ipp; English.
XX	
CC	The invention relates to modified inosine 5'-monophosphate dehydrogenase
CC	(IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo peptide domain
CC	substituted for a subdomain of a wild-type IMPDH polypeptide. The
CC	modified IMPDH polypeptides are useful for drug discovery, for
CC	therapeutic, diagnostic and prognostic procedures for detecting or
CC	quantifying modified IMPDH polypeptides and their corresponding nucleic
CC	acids. IMPDH polypeptides are also useful for generating antibodies, as
CC	diagnostic and prognostic markers of diseases, as targets for various
CC	therapeutic modalities, and to identify and isolate ligands and other
CC	agents that bind to modified IMPDH. These antibiotics may be used in
CC	diagnostic assays, imaging methodologies, therapeutic methods in the
CC	management of cancer or other proliferative-type diseases, and in
CC	purifying modified IMPDH polypeptides and for isolating related
CC	molecules such as wild type and mutant IMPDH polypeptides. The present
CC	sequence is human type II IMPDH-CSG variant cDNA.
XX	
SQ	Sequence 1155 BP; 277 A; 299 C; 312 G; 267 T; 0 other:
	Query Match            37.0%    Score 612.   DB 24:   Length 1155;
	Best Local Similarity   99.5%   Pred. No. 5.2e-282;
	Matches   812; Conservative %   0; Mismatches   4; Indels   0; Gaps
OY	777 CTGCCTGTGGAGCGACCATTTGCCATTCATAGATGACAAAGTATTAGCTTGACTTCTC 836
Db	340 CTGCTGTGTGGGCGACCATTTGCCATTCATAGATGACAAAGTATTAGCTTGACTTCTC 399
OY	837 GCCAGGCTGTGTGCATGTAGTGGTTTTGGACTTTTCCAGGAAATTCATCTTCAG 886
Db	400 GCCCAGGCTGTGTGCATGTAGTGGTTTTGGACTTTTCCAGGAAATTCATCTTCAG 459
OY	897 ATCATATGATCAAGTAGATCAAAGACAATAAACCCTTAATCCCAAGATTTGGAGGCAAT 956
Db	460 ATCATATGATCAAGTAGATCAAAGACAATAAACCCTTAATCCCAAGATTTGGAGGCAAT 519
OY	957 GTGTCTACTGCTGCCAGGCGCAAGAACCCTCAATTAATCAGAGTGTGATGCGCTTCAGGCTG 1016
Db	520 GTGTCTACTGCTGCCAGGCGCAAGAACCCTCAATTAATCAGAGTGTGATGCGCTTCAGGCTG 579
OY	1017 GGCAATGGGAATGGCTCCATCTGCATTTATCCAGGAATGCTGGCTGTGGGCGCCCCAA 1076
Db	560 GGCAATGGGAATGGCTCCATCTGCATTTATCCAGGAATGCTGGCTGTGGGCGCCCCAA 639
OY	1077 GCACAGCAGAGTACAGAGGTATAGATATGACAGGGCGCTTTGGTGTTCGATATGCT 1136
Db	640 GCACAGCAGAGTACAGAGGTATAGATATGACAGGGCGCTTTGGTGTTCGATATGCT 699
OY	1137 GATGAGGAATTCAAAATATGTGGGTTCATATTGCCAAAGCCTTGGCCCTTGGGCGCTCCACA 1196
Db	700 GATGAGGAATTCAAAATATGTGGGTTCATATTGCCAAAGCCTTGGCCCTTGGGCGCTCCACA 759
OY	1197 GTCATGATGGGCTCTCTCTCTGGCGCCACCACTGAGGCCCTTGATGAATCTTTTTC 1256
Db	760 GTCATGATGGGCTCTCTCTCTGGCGCCACCACTGAGGCCCTTGATGAATCTTTTTC 819
OY	1257 GATGGATCCGGGCTPAAGAAATATCGGGATATGGGTTCTCCATGCGCATGACAAGCAC 1316
Db	820 GATGGATCCGGGCTPAAGAAATATCGGGATATGGGTTCTCCATGCGCATGACAAGCAC 879
OY	1317 CTCAGCAGCCAGAACAGATATTCAGTGAAGCTGCACAAATTCAAAGTGGCGCCAGGAGTG 1376
Db	880 CTCAGCAGCCAGAACAGATATTCAGTGAAGCTGCACAAATTCAAAGTGGCGCCAGGAGTG 939

QY 1377 TCTGTCCTGTGCAGGACCAAGGCTCAATCCACAATTTGTCCCTTACCTGATTGCTGGC 1436  
|||||  
Db 940 TCTGTCCTGTGCAGGACCAAGGCTCAATCCACAATTTGTCCCTTACCTGATTGCTGGC 999  
QY 1437 ATCCAACTCATGCGCAGGACGATTTGGTCCAGAGCTTGGACCAAGTCCGAGCCATGATG 1436  
|||||  
Db 1000 ATCCAACTCATGCGCAGGACGATTTGGTCCAGAGCTTGGACCAAGTCCGAGCCATGATG 1059  
QY 1497 TACTCTGGGAGCTTAAATTGGAGAGAGAACTCTCAGCCCGAGTGAAGTGGCGTC 1556  
|||||  
Db 1060 TACTCTGGGAGCTTAAATTGGAGAGAGAACTCTCAGCCCGAGTGAAGTGGCGTC 1119  
QY 1557 CATAGCCTCATTTGCTATGAGAGAGCGCTTTTCTGA 1592  
|||||  
Db 1120 CATAGCCTCATTTGCTATGAGAGAGCGCTTTTCTGA 1155

RESULT 14  
AAD28928  
ID AAD28928 standard; cDNA; 1158 BP.  
XX  
AC AAD28928;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Human type II IMPDH-NSPL variant cDNA.  
XX  
KW Human: inosine 5'-monophosphate dehydrogenase; IMPDH: drug discovery;  
KM proliferative-type disease; cancer; imaging methodology; cytostatic;  
KW therapy: enzyme; variant; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1158  
FT /tag- a  
FT /product= "Human type II IMPDH-NSPL variant"  
XX  
PN WO200185952-A2.  
XX  
PD 15-NOV-2001.  
XX  
PE 10-MAY-2001; 2001MO-US15457.  
XX  
PR 10-MAY-2000; 2000US-203448P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Kyrstek SR, Sherif S, Witmer MR, Hollenbaugh DL, Yan N;  
PI Mourvieleff JE, Einspahr HM, Kish K;  
XX  
DR WPI; 2002-164105/21.  
DR P-PSDB; AAE18183.  
XX  
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
PS Claim 54; Fig 16; 161pp; English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the

CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is human type II IMPDH-NSPL variant cDNA.  
XX  
SQ Sequence 1158 BP; 279 A; 301 C; 309 G; 269 T; 0 other;  
Query Match 37.0%; Score 612; DB 24; Length 1158;  
Best Local Similarity 99.5%; Pred. No. 5.2e-282;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 777 CTGCTGTGTGGGCGCAGCCATTGGCAGCTCATGAGATGACAAGTATAGCGTGGACTTGGCTC 836  
|||||  
Db 343 CTGCTGTGTGGGCGCAGCCATTGGCAGCTCATGAGATGACAAGTATAGCGTGGACTTGGCTC 402  
QY 837 GCCCAGGCTGGTGGATGTAGTGGTTTGGACCTTCCGAGGAATTCATCTTCCAG 896  
|||||  
Db 403 GCCCAGGCTGGTGGATGTAGTGGTTTGGACCTTCCGAGGAATTCATCTTCCAG 462  
QY 897 ATCAATATGATCAAGTACATCAAGACAAATACCTTAATCTCCAAAGTCAATGGAGGCAAT 956  
|||||  
Db 463 ATCAATATGATCAAGTACATCAAGACAAATACCTTAATCTCCAAAGTCAATGGAGGCAAT 522  
QY 957 GTGGTCACTGTGCGCCAGGCCAAGACCTCATTTGATGACAGGTGTGATGTCCTGGCGGTG 1016  
|||||  
Db 523 GTGGTCACTGTGCGCCAGGCCAAGACCTCATTTGATGACAGGTGTGATGTCCTGGCGGTG 582  
QY 1017 GGCAATGGGAAGTGGCTCCATCTGCTCATTCACGGAATGTGCTGGCGTGGGCGGCCCAA 1076  
|||||  
Db 583 GGCAATGGGAAGTGGCTCCATCTGCTCATTCACGGAATGTGCTGGCGTGGGCGGCCCAA 642  
QY 1077 GCAACAGCAGTGTACAAGGTGTATGAGTATGACAGCGGCTTGTGTTCCGTCATTTGCT 1136  
|||||  
Db 643 GCAACAGCAGTGTACAAGGTGTATGAGTATGACAGCGGCTTGTGTTCCGTCATTTGCT 702  
QY 1137 GATGAGGAATCCAAATGTGGGTATTTGCCAAAGCCTTGGCCCTTGGGCGCTCCACA 1196  
|||||  
Db 703 GATGAGGAATCCAAATGTGGGTATTTGCCAAAGCCTTGGCCCTTGGGCGCTCCACA 762  
QY 1197 GTCATGATGGGCTCTCTCCGCTGGCTGCGCACCATGAGGCCCTGGTGAATCTCTTTTCC 1256  
|||||  
Db 763 GTCATGATGGGCTCTCTCCGCTGGCTGCGCACCATGAGGCCCTGGTGAATCTTTTCC 822  
QY 1257 GATGGATCCGGCTTAAAGAAATATCGCGGTATGGGTTCTCTCATGTCATGGACACAC 1316  
|||||  
Db 823 GATGGATCCGGCTTAAAGAAATATCGCGGTATGGGTTCTCTCATGTCATGGACACAC 882  
QY 1317 CTCAGCAGCCAGAAACAGATATTTTACGTAAGCTGACAAATTCAAAGTGGCCAGGAGTG 1376  
|||||  
Db 883 CTCAGCAGCCAGAAACAGATATTTTACGTAAGCTGACAAATTCAAAGTGGCCAGGAGTG 942  
QY 1377 TCTGTCCTGTGCAGGACCAAGGCTCAATCCACAATTTGTCCCTTACCTGATTGCTGGC 1436  
|||||  
Db 943 TCTGTCCTGTGCAGGACCAAGGCTCAATCCACAATTTGTCCCTTACCTGATTGCTGGC 1002  
QY 1437 ATCCAACTCATGCGCAGGACGATTTGGTCCAGAGCTTGGACCAAGTCCGAGCCATGATG 1496  
|||||  
Db 1003 ATCCAACTCATGCGCAGGACGATTTGGTCCAGAGCTTGGACCAAGTCCGAGCCATGATG 1062  
QY 1497 TACTCTGGGAGCTTAAATTGGAGAGAGAACTCTCAGCCCGAGTGAAGTGGCGTC 1556  
|||||  
Db 1063 TACTCTGGGAGCTTAAATTGGAGAGAGAACTCTCAGCCCGAGTGAAGTGGCGTC 1122  
QY 1557 CATAGCCTCATTTGCTATGAGAGAGCGCTTTTCTGA 1592  
|||||  
Db 1123 CATAGCCTCATTTGCTATGAGAGAGCGCTTTTCTGA 1158

RESULT 15  
AAF18367  
ID AAF18367 standard; DNA; 674 BP.  
XX  
AC AAF18367;

```
XX
DT 14-MAR-2001 (first entry)
DE Lung cancer associated polynucleotide sequence SEQ ID 386.
XX
KW Human: lung cancer associated protein; neuroprotective; cytoskeletal;
KW cardioactive; immunomodulatory; muscular active; vulnervary;
KW gastrointestinal; nephrotoxic; antineoplastic; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX
OS Homo sapiens.
XX
PN WO200055180-A2.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US05918.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
PI Ruben SM.
XX
DR WPI: 2000-587514/55.
XX P-PSDB: AAB58491.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX
PS Claim 1; Page 851-852; 1425pp; English.
XX
CC Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58498. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytoskeletal; cardioactive;
CC immunomodulatory; muscular active general; vulnervary; gastrointestinal
CC general; nephrotoxic; antineoplastic; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterization of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 674 BP; 179 A; 157 C; 186 G; 140 T; 12 other:
XX
Query Match 24.8%; Score 410; DB 21; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.6e-185;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DB 182 CCGAATTCAGGCCAATGAAGTTCGAAAGTGAGAAATATGACAGGATTCATCACA 241
|||||
QY 396 GACCTGTGTCCTTCAGCCCCAGAGATCGGTCGGATGTTTGTAGGCCAAGGCCGG 455
|||||
DB 242 GACCTGTGTCCTTCAGCCCCAGAGATCGGTCGGATGTTTGTAGGCCAAGGCCGG 301
|||||
QY 456 CATGTTTCTGCGGTATCCCAATCAGACAGAGCCGATGGGAGCCGCTTGTGGGC 515
|||||
DB 302 CATGTTTCTGCGGTATCCCAATCAGACAGAGCCGATGGGAGCCGCTTGTGGGC 361
|||||
QY 516 ATCATCTCCCTCCAGGACATGATTTCTCAAGAGAGAGAAATGACTG 565
|||||
DB 362 ATCATCTCCCTCCAGGACATGATTTCTCAAGAGAGAGAAATGACTG 411
|||||
RESULT 16
ABV44742
ID ABV44742 standard; cDNA; 607 BP.
XX
AC ABV44742:
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 44733.
XX
KW Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001MO-US05171.
XX
PR 17-FEB-2000; 2000US-18319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI: 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8871; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 607 BP; 141 A; 157 C; 171 G; 137 T; 1 other:
```

Query Match 24.4%: Score 404; DB 23; Length 607;  
Best Local Similarity 100.0%: Pred. No. 1,2e-182;  
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1101 GAGTATGACAGGGGCTTGTGTGTCGGTCATGTGTGATGAGAGAAATCCAAATGTGGGT 1160  
|||||  
DB 194 GAGTATGACAGGGGCTTGTGTGTCGGTCATGTGTGATGAGAGAAATCCAAATGTGGGT 253  
|||||  
QY 1161 CATATTGCCAAACCTTTGGCCCTTGGGCCCTCCACAGTATGATGGGCTCTCCGGCT 1220  
|||||  
DB 254 CATATTGCCAAACCTTTGGCCCTTGGGCCCTCCACAGTATGATGGGCTCTCCGGCT 313  
|||||  
QY 1221 GCCACACTGAGGCCCTGTGTGATATTCTTTCCGATGGGATCCGGCTAAAGAAATAT 1280  
|||||  
DB 314 GCCACACTGAGGCCCTGTGTGATATTCTTTCCGATGGGATCCGGCTAAAGAAATAT 373  
|||||  
QY 1281 CCGCGTATGGGTTCTCTCGATGCCATGAGACAGCACCCTCAGCAGCCAGAAAGATATTTTC 1340  
|||||  
DB 374 CCGCGTATGGGTTCTCTCGATGCCATGAGACAGCACCCTCAGCAGCCAGAAAGATATTTTC 433  
|||||  
QY 1341 AGTGAAGCTGACAAATTCAAAGTGGCCAGGAGTGTCTGTGTGCGAGCAAAAGG 1400  
|||||  
DB 434 AGTGAAGCTGACAAATTCAAAGTGGCCAGGAGTGTCTGTGTGCGAGCAAAAGG 493  
|||||  
QY 1401 TCATCCACAATTTTGTCTTACTGATTCGATTCGACATCCACACTCATCCAGACATT 1460  
|||||  
DB 494 TCATCCACAATTTTGTCTTACTGATTCGATTCGACATCCACACTCATCCAGACATT 553  
|||||  
QY 1461 GGTGCCAAGAGCTTGAACCCAGTCCGAGCCATGATGACTCTGG 1504  
|||||  
DB 554 GGTGCCAAGAGCTTGAACCCAGTCCGAGCCATGATGACTCTGG 597  
|||||

RESULT 17  
AAF94973  
ID AAF94973 standard; cDNA: 396 BP.  
XX  
AC AAF94973;  
XX

DT 23-MAY-2001 (first entry)

XX Human ovarian cancer associated coding sequence SEQ ID NO: 164.

XX Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.

XX Homo sapiens.

XX WO200118046-A2.

XX 15-MAR-2001.

XX 08-SEP-2000; 2000WO-US24827.

XX 10-SEP-1999; 99US-0394374.

XX 01-MAY-2000; 2000US-0561778.

XX 15-AUG-2000; 2000US-0640173.

XX 07-SEP-2000; 2000US-0656668.

XX (CORI-) CORIXA CORP.

XX Xu J, Stolk JA;

XX WPI; 2001-211395/21.

XX Isolated polypeptides associated with ovarian carcinomas, and the

XX nucleic acids that encode them, useful for the prevention diagnosis and

XX treatment of ovarian cancers -

XX Claim 18: Page 169; 189pp; English.

XX The present invention provides a number of coding sequences and proteins,

XX CC the over-expression of which is associated with ovarian carcinoma/cancer.

CC These can be used in the diagnosis, treatment and prevention of ovarian  
CC cancer, optionally by gene therapy or in the form of a vaccine. The  
CC present sequence is an example of one of these sequences.

XX Sequence 396 BP; 96 A; 119 C; 96 G; 85 T; 0 other;

Query Match 23.9%: Score 396; DB 22; Length 396;  
Best Local Similarity 100.0%: Pred. No. 8e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GACACGGCGGGTGTCTGTGTGGCCATGGCCGACTACCTGATTTAGTGGGGCAGCTCC 80  
|||||  
DB 1 GACACGGCGGGTGTCTGTGTGGCCATGGCCGACTACCTGATTTAGTGGGGCAGCTCC 60  
|||||  
QY 81 TACGTGCGAGACGAGGACTCAGACAGCAGCTCTTCAACTGGCGAGACGGCTCACC 140  
|||||  
DB 61 TACGTGCGAGACGAGGACTCAGACAGCAGCTCTTCAACTGGCGAGACGGCTCACC 120  
|||||  
QY 141 TACATGACTTTTCTCATTTCTCCTGGGTACATGACCTTCACTCAGACAGGTGACCTG 200  
|||||  
DB 121 TACATGACTTTTCTCATTTCTCCTGGGTACATGACCTTCACTCAGACAGGTGACCTG 180  
|||||  
QY 201 ACTTTGCTCTGACCAAGAAATCACTTTAAGACCCCACTGTTTCTCTCCATGAC 260  
|||||  
DB 181 ACTTTGCTCTGACCAAGAAATCACTTTAAGACCCCACTGTTTCTCTCCATGAC 240  
|||||  
QY 261 ACATGACAGAGGCTGGATGGCCATGACAAATGGCGCTTACAGCGGTAATTGGCTCATC 320  
|||||  
DB 241 ACATGACAGAGGCTGGATGGCCATGACAAATGGCGCTTACAGCGGTAATTGGCTCATC 300  
|||||  
QY 321 CACCACACTGTACACCTGATTCAGGCCAATGAAGTTGGAAGTGAAGAAATATGAA 380  
|||||  
DB 301 CACCACACTGTACACCTGATTCAGGCCAATGAAGTTGGAAGTGAAGAAATATGAA 360  
|||||  
QY 381 CAGGATTCATCAGACACCTGTGTCTCAGCCCC 416  
|||||  
DB 361 CAGGATTCATCAGACACCTGTGTCTCAGCCCC 396  
|||||

RESULT 18  
ABT03240  
ID ABT03240 standard; cDNA: 396 BP.  
XX  
AC ABT03240;  
XX

DT 05-SEP-2002 (first entry)

XX Human ovarian carcinoma associated coding sequence SEQ ID NO: 164.

XX Human, ovarian cancer; ovarian carcinoma; gene therapy; immunotherapy;

XX cytostatic; gene; ss.

XX Homo sapiens.

XX WO200239885-A2.

XX 23-MAY-2002.

XX 13-NOV-2001; 2001WO-US45395.

XX 14-NOV-2000; 2000US-0713550.

XX 03-APR-2001; 2001US-0825294.

XX 02-OCT-2001; 2001US-0970966.

XX (CORI-) CORIXA CORP.

XX Xu J, Stolk JA, Algate PA, Fling SP, Molesh DA;

XX WPI; 2002-500186/53.

XX Novel ovarian cancer polypeptide and polynucleotide, useful for

XX PT detecting the presence of ovarian cancer in a patient, and in

XX pharmaceutical compositions, e.g. vaccines, for treating ovarian cancer



PT -  
XX  
PS Example 1: Page 165; 197pp; English.  
XX

CC The present invention provides human ovarian cancer associated proteins  
CC and coding sequences. The sequences can be used in the diagnosis and  
CC treatment of ovarian cancers. The present sequence is a coding sequence  
CC of the invention.  
XX  
SQ Sequence 396 BP; 96 A; 119 C; 96 G; 85 T; 0 other;

Query Match 23.9%; Score 396; DB 24; Length 396;  
Best Local Similarity 100.0%; Pred. No. 8e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GACACGGCGGCTGCTCTGTTGGCCATGGCCGACTACCTGATTAGTGGGGCAGCTCC 80  
DB 1 GACACGGCGGCTGCTCTGTTGGCCATGGCCGACTACCTGATTAGTGGGGCAGCTCC 60  
QY 81 TAGCTCCAGACGAGGACTCAGACAGACGACTCTTCACTGACGGGACGGCCCTCACC 140  
DB 61 TAGCTCCAGACGAGGACTCAGACAGACGACTCTTCACTGACGGGACGGCCCTCACC 120  
QY 141 TACATGACTTTTCTCATTTCTCCCTGGTACATGACTTCACTGACAGCAGGTGACCTG 200  
DB 121 TACAATGACTTTTCTCATTTCTCCCTGGTACATGACTTCACTGACAGCAGGTGACCTG 180  
QY 201 ACTTGTGCTGTGACCAAGAAATCACTTTAAGCCCGACTGGTTCCCTCCCATGGAC 260  
DB 181 ACTTGTGCTGTGACCAAGAAATCACTTTAAGCCCGACTGGTTCCCTCCCATGGAC 240  
QY 261 ACAGTCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGCTTTGGCTTCATC 320  
DB 241 ACAGTCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGCTTTGGCTTCATC 300  
QY 321 CACCACAACGTGACACCTGAATTCAGGCCCAATGAAGTTCCGGAAGTGAAGAATATGAA 380  
DB 301 CACCACAACGTGACACCTGAATTCAGGCCCAATGAAGTTCCGGAAGTGAAGAATATGAA 360  
QY 381 CAGGATTTCATCAGACACCTGTGCTCCAGCCCC 416  
DB 361 CAGGATTTCATCAGACACCTGTGCTCCAGCCCC 396

RESULT 19

ABL48923  
ID ABL48923 standard; cDNA: 396 BP.

XX ABL48923;

XX 18-JUN-2002 (first entry)

XX Ovarian carcinoma sequence isolate 25775.

XX Ovarian cancer; cancer therapy; vaccine; gene therapy; tumour; cancer;

XX SS.

XX Homo sapiens.

XX US2002004491-A1.

XX 10-JAN-2002.

XX 03-APR-2001: 2001US-0825294.

XX 10-SEP-1999: 99US-0394374.

XX 01-MAY-2000: 2000US-0561778.

XX 15-AUG-2000: 2000US-0640173.

XX 07-SEP-2000: 2000US-0656668.

XX 14-NOV-2000: 2000US-0713550.

XX (XUJ/) XU J.

XX PA (STOL/) STOLK J A.

PA (ALGA/) ALGATE P A.  
PA (FLIN/) FLING S P.  
XX

PI Xu J, Stolk JA, Algate PA, Fling SP;

XX WPI: 2002-171027/22.

XX Ovarian tumour polypeptide and polynucleotide useful in diagnosis,

XX prevention and/or treatment of cancer, especially ovarian cancer.

PS Example 1: Page 93; 131pp; English.

CC The invention relates to ovarian tumour polynucleotides and polypeptides  
CC that may be utilised in cancer therapy, for example in a vaccine or  
CC gene therapy. Polypeptides and polynucleotides of the invention are  
CC useful for detecting a cancer in a patient, for stimulating and/or  
CC expanding T-cells specific for a tumour protein, and for inhibiting the  
CC development of a cancer in a patient. They are also useful for  
CC stimulating an immune response in a patient, and for treating a cancer in  
CC a patient and for determining the presence of a cancer in a patient.  
CC The isolated polynucleotides of the invention are useful for their  
CC ability to selectively form duplex molecules with complementary stretches  
CC of the entire desired gene or gene fragments, and for designing and  
CC preparing ribozyme molecules for inhibiting expression of tumour  
CC polypeptides in tumour cells. Polypeptides and polynucleotides of the  
CC invention are also useful in recombinant DNA molecules to direct  
CC expression of a polypeptide in appropriate host cells. The sequences  
CC given in records ABL48760-ABL48956 represent polynucleotides encoding  
CC ovarian carcinoma proteins.

SQ Sequence 396 BP; 96 A; 119 C; 96 G; 85 T; 0 other;

Query Match 23.9%; Score 396; DB 24; Length 396;  
Best Local Similarity 100.0%; Pred. No. 8e-179;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GACACGGCGGCTGCTCTGTTGGCCATGGCCGACTACCTGATTAGTGGGGCAGCTCC 80  
DB 1 GACACGGCGGCTGCTCTGTTGGCCATGGCCGACTACCTGATTAGTGGGGCAGCTCC 60  
QY 81 TAGCTCCAGACGAGGACTCAGACAGACGACTCTTCACTGACGGGACGGCCCTCACC 140  
DB 61 TAGCTCCAGACGAGGACTCAGACAGACGACTCTTCACTGACGGGACGGCCCTCACC 120  
QY 141 TACATGACTTTTCTCATTTCTCCCTGGTACATGACTTCACTGACAGCAGGTGACCTG 200  
DB 121 TACAATGACTTTTCTCATTTCTCCCTGGTACATGACTTCACTGACAGCAGGTGACCTG 180  
QY 201 ACTTGTGCTGTGACCAAGAAATCACTTTAAGCCCGACTGGTTCCCTCCCATGGAC 260  
DB 181 ACTTGTGCTGTGACCAAGAAATCACTTTAAGCCCGACTGGTTCCCTCCCATGGAC 240  
QY 261 ACAGTCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGCTTTGGCTTCATC 320  
DB 241 ACAGTCACAGAGGCTGGATGGCCATAGCAATGGCCCTTACAGCGGCTTTGGCTTCATC 300  
QY 321 CACCACAACGTGACACCTGAATTCAGGCCCAATGAAGTTCCGGAAGTGAAGAATATGAA 380  
DB 301 CACCACAACGTGACACCTGAATTCAGGCCCAATGAAGTTCCGGAAGTGAAGAATATGAA 360  
QY 381 CAGGATTTCATCAGACACCTGTGCTCCAGCCCC 416  
DB 361 CAGGATTTCATCAGACACCTGTGCTCCAGCCCC 396

RESULT 20

AAC03976  
ID AAC03976 standard; cDNA: 554 BP.

XX AAC03976;

XX 06-OCT-2000 (first entry)

XX

```

DE Human secreted protein 5' EST, SEQ ID NO: 3974.
XX
XX Human: 5' EST: expressed sequence tag: secreted protein; cDNA isolation:
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
XX EP1033401-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
XX
XX P-PSDB: AAG03970.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 3974; 71bp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX
XX Sequence 554 BP; 122 A; 159 C; 147 G; 118 T; 8 other;
XX
XX
XX Query Match 21.8%; Score 361; DB 21; Length 554;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-162;
XX Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 9 GCGGTCCTGGAGACACGCGGCTGCTGTTGGCCATGGCGACTGATTAGT 68
XX |||||||
XX 25 GCGGTCCTGGAGACACGCGGCTGCTGTTGGCCATGGCGACTGATTAGT 84
XX
XX 69 GGGGGACCTCTAGCTGCCAGACGAGACTACAGACAGCAGCTCTTCAACTGCGGA 128
XX |||||||
XX 85 GGGGGACCTCTAGCTGCCAGACGAGACTACAGACAGCAGCTCTTCAACTGCGGA 144
XX
XX 129 GACGGCTCAGCAATGACTTTCATTTCCTGGGAGCTGACTGACTTCACTGAGAC 188
XX |||||||
XX 145 GACGGCTCAGCAATGACTTTCATTTCCTGGGAGCTGACTGACTTCACTGAGAC 204
XX
XX 189 CAGGTGACCTGACTTCTGCTGACCAAGAAATCACTTAAAGACCCCACTGGTTTTC 248
XX |||||||
XX 205 CAGGTGACCTGACTTCTGCTGACCAAGAAATCACTTAAAGACCCCACTGGTTTTC 264
XX
XX 249 TCTCCCATGACACAGTACAGAGGCTGGAGTGGCCATAGCAATGGCGCTTACAGGCGGT 308
XX |||||||
XX 265 TCTCCCATGACACAGTACAGAGGCTGGAGTGGCCATAGCAATGGCGCTTACAGGCGGT 324
XX
XX 309 ATTGGTTTATCCACCAACCTTACACTGATTCAGAGCCCAATGAAGTGGGAAGTG 368
XX |||||||
XX 325 ATTGGTTTATCCACCAACCTTACACTGATTCAGAGCCCAATGAAGTGGGAAGTG 384
XX

```

```

QY 369 A 369
DB 385 A 385
XX
XX RESULT 21
XX ABV37312/c
XX ID ABV37312 standard; cDNA; 420 BP.
XX
XX ABV37312;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 37303.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
XX 16-MAR-2000; 2000US-189862P.
XX
XX 25-MAY-2000; 2000US-207454P.
XX
XX 09-JUN-2000; 2000US-211314P.
XX
XX 18-JUL-2000; 2000US-219007P.
XX
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI: 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 7666; 11750bp; English.
XX
XX
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 420 BP; 87 A; 120 C; 105 G; 108 T; 0 other;
XX
XX
XX Query Match 17.9%; Score 296; DB 23; Length 420;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-131;
XX Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 619 GCATCAGCTGGAAGGAGCAATGAATTCGTGAGCGCAGCAAGAAGGAAGTTGCCCA 678
XX |||||||
XX DB 398 GCATCAGCTGGAAGGAGCAATGAATTCGTGAGCGCAGCAAGAAGGAAGTTGCCCA 339
XX |||||||
XX QY 679 TTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCCGACAGACTGAAGAGAAATC 738
XX |||||||

```

Db 338 TTGTAATGAAGATGATGAGCTTGTGGCCATTCATGCGCGAGACGCTGAAGAAGATC 279  
QY 739 GGGACTACCCACTAGCCTCCCAAGATGCCAAGAACAGCTGCTGTGTGGGACGCCATTG 798  
Db 278 GGGACTACCCACTAGCCTCCCAAGATGCCAAGAACAGCTGCTGTGTGGGACGCCATTG 219  
QY 799 GCACATCATGAGATGACAACTATAGCTGAGCTTGTGCTGCGCCAGAGTGTGTGATGTAG 858  
Db 218 GCACATCATGAGATGACAACTATAGCTGAGCTTGTGCTGCGCCAGAGTGTGTGATGTAG 159  
QY 859 TGGTTTTGGAAGCTTCCACAGGAATTCCATCTTCCAGATCAATATGCTCAAGTAC 914  
Db 158 TGGTTTTGGAAGCTTCCACAGGAATTCCATCTTCCAGATCAATATGCTCAAGTAC 103  
RESULT 22  
AAZ80760 standard: cDNA: 675 BP.  
ID AAZ80760  
XX AAZ80760:  
AC  
XX  
DT 07-APR-2000 (first entry)  
XX  
DE Human colon cancer cell line SW480 cDNA clone SEQ ID: NO: 844.  
XX  
KW Human: gene expression product; diagnosis: tumour; colon cancer;  
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;  
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;  
KW hyperplasia; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO9964576-A2.  
PD 16-DEC-1999.  
XX  
PF 09-JUN-1999: 99MO-IB01062.  
XX  
PR 10-JUN-1998: 98US-0088801.  
XX  
PA (FARB ) BAYER CORP.  
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;  
PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;  
PI Schlegel R;  
XX  
DR WPI: 2000-087220/07.  
XX  
PT Novel nucleic acids, used to develop products for the diagnosis and  
PT treatment of disorders involving unwanted cell proliferation,  
PT particularly cancers, especially colon cancer  
XX  
PS Claim 15: Page 466-467: 469pp: English.  
XX  
CC AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from  
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The  
CC cDNA clones can be used to generate antisense oligonucleotides which  
CC can be used for antisense therapy. Methods and products from the present  
CC invention can be used for identifying and/or classifying cancerous cells  
CC present in a human tumour, particularly in solid tumors, e.g.,  
CC carcinomas and sarcomas, e.g., breast or colon cancers. The cDNA clones  
CC can be used for developing agents for the diagnosis and treatment of  
CC disorders involving unwanted cell proliferation, such as neoplasia,  
CC dysplasia or hyperplasia.  
XX  
SQ Sequence 675 BP: 169 A; 150 C; 186 G; 161 T; 9 other:  
Query Match 17.1%; Score 283; DB 21: Length 675;  
Best Local Similarity 100.0%; Pred. No. 7.6e-125;  
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 331 GTACACCTGATTCACAGCCCAATGAAGTTGGAAAGTGAAGAATATGACAGGATTCA 390  
|||||

Db 2 GTACACCTGATTCACAGCCCAATGAAGTTGGAAAGTGAAGAATATGACAGGATTCA 61  
QY 391 TCACAGACCCCTGTGCTCTTACGCCCAAGATCGCGGATGTTTTTGAGGCCAAG 450  
Db 62 TCACAGACCCCTGTGCTCTTACGCCCAAGATCGCGGATGTTTTTGAGGCCAAG 121  
QY 451 CCCGGCATGCTTCTGCGGTATCCCATTCACAGACAGCCCGGATGGGAGCCCTTGG 510  
Db 122 CCCGGCATGCTTCTGCGGTATCCCATTCACAGACAGCCCGGATGGGAGCCCTTGG 181  
QY 511 TGGGCATCATCTCTCCACAGGACATTTGTTCTCAAAAGAGGAGACATGATGTTCT 570  
Db 182 TGGGCATCATCTCTCCACAGGACATTTGTTCTCAAAAGAGGAGACATGATGTTCT 241  
QY 571 TGGAGAGATTAATGACAAAGAGGAGACATGTTGTTAGCCCC 613  
Db 242 TGGAGAGATTAATGACAAAGAGGAGACATGTTGTTAGCCCC 284  
RESULT 23  
ABV07378/c  
ID ABV07378 standard: cDNA: 391 BP.  
XX  
AC ABV07378:  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 7369.  
XX  
KW Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001: 2001MO-US05171.  
XX  
PR 17-FEB-2000: 2000US-183319P.  
PR 16-MAR-2000: 2000US-189862P.  
PR 25-MAY-2000: 2000US-207454P.  
PR 09-JUN-2000: 2000US-211314P.  
PR 18-JUL-2000: 2000US-219007P.  
PR 13-DEC-2000: 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JE;  
PI  
XX  
DR WPI: 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1: Page 1192: 11750pp: English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (i) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement, (ii) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient.

CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 391 BP; 81 A; 107 C; 96 G; 102 T; 5 other;

Query Match	16.1%;	Score 266;	DB 23;	Length 391;
Best Local Similarity	100.0%;	Pred. No. 1e-116;		
Matches 266;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.0;

Qy 649 TGCAGCGCAGCAAGAAGGAAAGTTGCCATTGAATGAGATGATGACCTTGTGGCCA 708  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 328 TGCAGCGCAGCAAGAAGGAAAGTTGCCATTGAATGAGATGATGACCTTGTGGCCA 269

Qy	709	TCAATGCCCCGACAGACTGAAGAAGAA	TCGGCACTACCCACTCAAGAATGCCA	768
Db	268	TCATTGCCCGACAGACTGAAGAAGAA <td>TCGGCACTACCCACTAGCCTCCAAAGATGCCA</td> <td>209</td>	TCGGCACTACCCACTAGCCTCCAAAGATGCCA	209

QY 769 AGAACAGCTGCTGTGTGGGCGACCAATTGGCACTCATGAGATGACAACTATAGCGTGG 828  
|||||  
Db 208 AGAAACAGCTGCTGTGTGGGCGACCAATTGGCACTCATGAGATGACAACTATAGCGTGG 1499

QY 829 ACTTCTCGCCCGACGCTGTTGATGTAGTGTGTTTGACCTCTTCCACGGAATTCCA 888  
|||||  
Db 148 ACTTGTCCCGCAGCGCTGTGTGTGATGTAGTGTGTTTGACTCTTCCACGGAATTCCA 89  
|||||

Qy	889	TCTTCAGATCAATATGATCAAGTAC	914
Db	88	TCTTCAGATCAATATGATCAAGTAC	63

RESULT 24  
ABO59475/C

ID	ABQ59475	Scandalo; CDNA; 5' / BP.
XX		
AC	ABQ59475;	
XX		

DI	02-AUG-2002 (first entry)
XX	
DE	Human colon cancer related nucleotide sequence SEQ ID NO:3170.
XX	

OS Homo sapiens  
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX

XX	WO200229086-A2.
PN	
XX	
PD	11-APR-2002

XX	02-OCT-2001; 2001WO-US30732.
PF	
XX	
DB	02-OCT-2000; 2000US-327271D

XX (FARB ) BAYER CORP.  
PA  
XX

PI Thiagalingam A, Lewis ME;  
XX  
DR WPI: 2002-42615/45.  
...

PT New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell  
PT or tissue type, and in antisense therapy -

PS	Claim 1; Fig 1; 796pp; English.
XX	
CC	ABQ56306 to ABQ60787 represent

ABO56310 to ABO60787 represent isolated nucleic acids (1) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABOC07776 to ABOC07787 nucleic acid sequences. (1) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (1) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridises to (1) in a cell. A probe/primer derived from (1) can be used for determining the presence of a nucleic acid which

CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence or  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (II) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists.  
XX  
XX Sequence 577 BP; 136 A; 135 C; 119 G; 167 T; 20 other:

Query Match	16.0 %	Score 264	DB 24	Length 577
Best Local Similarity	99.7 %	Pred. No. 9e-116		
Matches 314; Conservative	0	Mismatches	1	Indels 0
				Gaps 0

QY 1185 GGGGCGCTCCACAGTCATGATGGGCTCTCCTGGTGCACCACCTGA6GCCCTGTGTAA 1244  
|||||  
Db 316 GGGGCTCCACAGTCATGATGGGCTCTCCTGGTGCACCACCTGA6GCCCTGTGTAA 257

QY 1245 TACTTCTTTTCGATGGATCCGGCTAAAGAAATATTCGGGTATGGGTCTCTCGATGCC 1304  
|||||  
Db 256 TACTTCTTTTCGATGGATCCGGCTAAAGAAATATTCGGGTATGGGTCTCTCGATGCC 197

QY 1305 ATGCACAAGCACCCTCAGCAGCCAGAAGATATTTCAGTGAAGCTGACAAATCAAACTG 1364  
|||||  
Db 196 ATGCACAAGCACCCTCAGCAGCCAGAAGATATTTCAGTGAAGCTGACAAATCAAACTG 137

Qy 1365 GCCCAGGAGTGTCTGGTCTGTGCAGGACAAAGGTCATTCACAAATTTGTCCTTAC 1424

Db 136 GCCCAGGAGTGTCTGGTCTGTGCAGGACAAAGGTCATTCACAAATTTGTCCTTAC 77

Oy 1425 CTGATTCGTCGGCATCCACACACTCATTCGCCAGACATTGGTGCCAGAGCTTGACCCCAAGTC 1484  
|||  
Db 76 CTGATTCGTCGGCATCCACACACTCATTCGCCAGACATTGGTGCCAGAGCTTGACCCCAAGTC 17

QY 1485 CGAGCCATGATGTAC 1499  
 |||||  
 Db 16 CGAGCCATGATGTAC 2

RESULT 25  
ARN95699

ID	ABN95699	standard; DNA; 6193 BP.
XX		
AC	ABN95699;	
XY		

DT	13-AUG-2002 (first entry)
XX	
DE	Gene #2197 used to diagnose liver cancer
XY	

KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KM metastatic liver tumour; cystostatic; expression profile; disease state;  
KM disease progression; drug toxicity; drug efficacy; drug metabolism.

OS	Homo sapiens.
XX	
PN	WO200229103-A2.

PD	11-APR-2002.
XX	
PF	02-OCT-2001; 2001WO-US30589.

PR 02-OCT-2000; 2000US-237054P.  
XX  
PA (GENE-) GENE LOGIC INC.

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX  
DR WPI; 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer,  
PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample -  
XX  
XX Claim 1: SEQ ID NO 2197: 298pp: English.  
XX  
CC The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumor in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SO Sequence 6193 BP: 1404 A: 1597 C: 1628 G: 1564 T: 0 other:  
Query Match 14.6%: Score 242: DB 24: Length 6193:  
Best Local Similarity 100.0%: Pred. No. 2.8e-105:  
Matches 242: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
YY 1101 GAATATGACAGCGCCCTTGTGTCCTCCGTCATTTGCTGATGAGGAATCAAAATGTGGGT 1160  
DB 5210 GAGTATGACAGCGCGCTTGTGTCCTCCGTCATTTGCTGATGAGGAATCAAAATGTGGGT 5269  
YY 1161 CATATTTGCCAAGCCCTTGGCCCTTGGGCCCTCCACAGTCATGATGAGGCTCTCTCTGGCT 1220  
DB 5270 CATATTTGCCAAGCCCTTGGCCCTTGGGCCCTCCACAGTCATGATGAGGCTCTCTCTGGCT 5329  
YY 1221 GCCACCACTGAGCGCCCTGTGTCATCTTCTTCCGATGGATCCGCTAAAGAAATAT 1280  
DB 5330 GCCACCACTGAGCGCCCTGTGTCATCTTCTTCCGATGGATCCGCTAAAGAAATAT 5389  
YY 1281 CGCGGATAGGTTCTCTCGATGCCATGACAGACACCTTCAGCAGCCAGACAGATATTTTC 1340  
DB 5390 CGCGGATAGGTTCTCTCGATGCCATGACAGACACCTTCAGCAGCCAGACAGATATTTTC 5449  
YY 1341 AG 1342  
DB 5450 AG 5451  
RESULT 26  
ABO58310  
ID ABO58310 standard: cDNA: 558 BP.  
XX  
XX ABO58310:  
XX  
DT 02-AUG-2002 (first entry)  
XX  
DE Human colon cancer related nucleotide sequence SEQ ID NO:2005.  
XX  
XX Human: colon cancer: cancer: tissue profiling: forensic  
XX genetic analysis: diagnostic: antisense therapy: gene  
XX ss.  
OS Homo sapiens.  
XX  
XX WO200229086-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001: 2001WO-US30732.  
XX  
XX 02-OCT-2000: 2000US-237271P.  
XX  
XX (FARB ) BAYER CORP.  
XX  
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA:

PI Thielglingam A, Lewis ME:  
XX  
XX WPI: 2002-426115/45.  
XX  
PT New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell  
PT or tissue type, and in antisense therapy -  
XX  
XX Claim 1: Fig 1: 796pp: English.  
XX  
CC ABO56306 to ABO60787 represent isolated nucleic acids (I) differentially  
CC expressed in cancer tissues. ABO78993 to ABO79004 represent proteins  
CC encoded by the ABO60776 to ABO60787 nucleic acid sequences. (I) can be  
CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC microarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists.  
XX  
SO Sequence 558 BP: 151 A: 119 C: 143 G: 123 T: 22 other:  
Query Match 13.9%: Score 230: DB 24: Length 558:  
Best Local Similarity 100.0%: Pred. No. 1.6e-99:  
Matches 230: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
YY 343 TCCAGGCGAATGAAGTTGGGAAGTGAAGAAATATACAGGATTCATCAGACCTTG 402  
DB 11 TCCAGGCGAATGAAGTTGGGAAGTGAAGAAATATATACAGGATTCATCAGACCTTG 70  
YY 403 TGGTCTCTCAGCCCCCAAGATCGCTCGGAGTGTCTTTTTCAGGCGCAAGCCCGCATGTT 462  
DB 71 TGGTCTCTCAGCCCCCAAGATCGCTCGGAGTGTCTTTTTCAGGCGCAAGCCCGCATGTT 130  
YY 463 TCTGCGGTATCCCAATCACAGACAGCGCGGATGGGAGCCCTTGTTGGGATCATCT 522  
DB 131 TCTGCGGTATCCCAATCACAGACAGCGCGGATGGGAGCCCTTGTTGGGATCATCT 190  
YY 523 CCTCCAGGCAATGATTTTTCCTCAAGAGAGGAGACATGACTGTTCTTG 572  
DB 191 CCTCCAGGCAATGATTTTTCCTCAAGAGAGGAGACATGACTGTTCTTG 240  
RESULT 27  
AAS18240  
ID AAS18240 standard: DNA: 10557 BP.  
XX  
XX AAS18240:  
XX  
DT 25-FEB-2002 (first entry)  
XX  
DE Reference sequence for human IMPDH2 gene showing SNPs.  
XX  
XX Human: single nucleotide polymorphism: SNP: IMPDH2: chromosome 3p21.2;  
XX IMP dehydrogenase 2; haplotyping; genotyping; cancer; cytostatic; ds.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX  
XX Key location/Qualifiers  
XX replace (1564, C)  
XX  
XX /tag- a  
XX /standard name= "Single nucleotide polymorphism"  
XX /note= "Polymorphic site 1 (PS1)"  
XX replace (11758, T)  
XX /tag- b  
XX /standard name= "Single nucleotide polymorphism"

```

FT      /note= "Polymorphic site 2 (PS2)"
FT      replace (1810, T)
FT      /tag= c
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 3 (PS3)"
FT      replace (1902, C)
FT      /tag= d
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 4 (PS4)"
FT      1972..2069
FT      /tag= e
FT      /number= 1
FT      1972..6937
FT      /tag= f
FT      /product= "IMPDH2"
FT      /note= "Coding region specifically claimed in
FT      claim 24"
FT      2070..2514
FT      /tag= g
FT      /number= 1
FT      replace (2511, A)
FT      /tag= h
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 5 (PS5)"
FT      2515..2563
FT      /tag= i
FT      /number= 2
FT      2564..2788
FT      /tag= j
FT      /number= 2
FT      replace (2628, G)
FT      /tag= k
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 6 (PS6)"
FT      2789..2890
FT      /tag= l
FT      /number= 3
FT      2891..2997
FT      /tag= m
FT      /number= 3
FT      2998..3072
FT      /tag= n
FT      /number= 4
FT      3073..3399
FT      /tag= o
FT      /number= 4
FT      replace (3294, C)
FT      /tag= p
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 7 (PS7)"
FT      3400..3606
FT      /tag= q
FT      /number= 5
FT      replace (3459, C)
FT      /tag= r
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 8 (PS8)"
FT      3607..4263
FT      /tag= s
FT      /number= 5
FT      replace (4168, A)
FT      /tag= t
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 9 (PS9)"
FT      replace (4252, T)
FT      /tag= u
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 10 (PS10)"
FT      4264..4351
FT      /tag= v
FT      /number= 6
FT      4352..4424
FT      /tag= w

FT      /number= 6
FT      4425..4624
FT      /tag= x
FT      /number= 7
FT      4625..4701
FT      /tag= y
FT      /number= 7
FT      replace (4634, C)
FT      /tag= z
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 11 (PS11)"
FT      4702..4792
FT      /tag= aa
FT      /number= 8
FT      4793..4891
FT      /tag= ab
FT      /number= 8
FT      4892..4987
FT      /tag= ac
FT      /number= 9
FT      4988..6052
FT      /tag= ad
FT      /number= 9
FT      6053..6196
FT      /tag= ae
FT      /number= 10
FT      6197..6279
FT      /tag= af
FT      /number= 10
FT      6280..6424
FT      /tag= ag
FT      /number= 11
FT      6425..6518
FT      /tag= ah
FT      /number= 11
FT      replace (6495, T)
FT      /tag= ai
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 12 (PS12)"
FT      6519..6662
FT      /tag= aj
FT      /number= 12
FT      6663..6742
FT      /tag= ak
FT      /number= 12
FT      6743..6826
FT      /tag= al
FT      /number= 13
FT      6827..6915
FT      /tag= am
FT      /number= 13
FT      6916..6937
FT      /tag= an
FT      /number= 14
FT      replace (7088, T)
FT      /tag= ao
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "Polymorphic site 13 (PS13)"

FT      WO200177363-A2.
FT      PN
FT      18-OCT-2001.
FT      PD
FT      XX
FT      11-APR-2001; 2001WO-US11851.
FT      PR
FT      11-APR-2000; 2000US-196248P.
FT      PA
FT      (GENA-) GENAISSANCE PHARM INC.
FT      XX
FT      Chew A, Choi JY, Koshy B, Lee HH, Stephens JC;
FT      PI
FT      WPI: 2002-041297/05.
FT      DR
FT      P-PSDB; AAU10695.

```

XX New isolated polynucleotide having polymorphic variant of IMP2  
PT dehydrogenase gene, useful for studying expression of the gene in vivo,  
PT and for testing efficacy of therapeutic agents for cancer in biological  
PT system -  
XX  
PS Claim 19; Fig 1; 70pp; English.  
XX  
CC The present invention relates to novel single nucleotide polymorphisms  
CC (SNPs) in the human IMP dehydrogenase 2 (IMPDH2) gene located on  
CC chromosome 3p21.2, and methods for haplotyping and/or genotyping the  
CC IMPDH2 gene in an individual. The methods of the invention make use of  
CC allele-specific oligonucleotides (ASOs) as probes and primers and/or  
CC primer-extensions oligonucleotides for detecting the IMPDH2 gene  
CC polymorphisms. The polynucleotides and screened compounds are  
CC useful for (developing) treatment of diseases associated with IMPDH2  
CC activity, such as cancer. The present sequence represents a reference  
CC sequence for the IMPDH2 gene which shows the variations in the gene.  
XX  
SQ Sequence 10557 BP; 2348 A; 2852 C; 2963 G; 2394 T; 0 other:  
XX  
Query Match 12.6%; Score 209; DB 24; Length 10557;  
Best Local Similarity 100.0%; Pred. No. 1.6e-89;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 370 AGAATATGAACAGGGATTCATCAGACCCCTGTCCTCAGCCCGCAGATGCGCTGC 429  
DB 3398 AGAATATGAACAGGGATTCATCAGACCCCTGTCCTCAGCCCGCAGATGCGCTGC 3457  
OY 430 GGCATGTTTGTAGCCCAAGCCCGCATGTTCTCGCGTATCCCA TCACAGACACAG 489  
DB 3458 GGCATGTTTGTAGCCCAAGCCCGCATGTTCTCGCGTATCCCA TCACAGACACAG 3517  
OY 490 GCCGATGGGAGCGCGTTGTGGGCAATCTCTCCAGGACATTTATTTTCTCAAG 549  
DB 3518 GCCGATGGGAGCGCGTTGTGGGCAATCTCTCCAGGACATTTATTTTCTCAAG 3577  
OY 550 AGGAGAACATGACTGTTTCTTGGAGAG 578  
DB 3578 AGGAGAACATGACTGTTTCTTGGAGAG 3606  
RESULT 28  
ABV14868  
ID ABV14868 standard; cDNA; 379 BP.  
XX  
AC ABV14868;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 14859.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;

XX  
DR WPI: 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 2488; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (1) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (1) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 379 BP; 87 A; 93 C; 108 G; 91 T; 0 other:  
XX  
Query Match 9.3%; Score 154; DB 23; Length 379;  
Best Local Similarity 100.0%; Pred. No. 3.2e-63;  
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1101 GAGTATGACAGCGCGCTTGGTTCGCGTCAATTCATGAGGAATCCAAATGTGGT 1150  
DB 225 GAGTATGACAGCGCGCTTGGTTCGCGTCAATTCATGAGGAATCCAAATGTGGT 284  
OY 1161 CATATGCGAAAGCCTTGCCCTTGAGGCGCTCCACAGTCATGATGGCTCTCTGCGT 1220  
DB 285 CATATGCGAAAGCCTTGCCCTTGAGGCGCTCCACAGTCATGATGGCTCTCTGCGT 344  
OY 1221 GCCACACCTGAGCGCCCTGGTGGAATCTCTTTT 1254  
DB 345 GCCACACCTGAGCGCCCTGGTGGAATCTCTTTT 378  
RESULT 29  
ABV38955  
ID ABV38955 standard; cDNA; 437 BP.  
XX  
AC ABV38955;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 38946.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI

```
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 7915; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 437 BP; 104 A; 114 C; 123 G; 96 T; 0 other:
XX
Query Match 9.3%; Score 153; DB 23; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.6e-63;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1101 GAGTATGACAGGCGCTTGGTGTCCGGTCATTGCTGATGAGAGAAATCAAAATGTGGGT 1160
DB 285 GAGATGACAGGCGCTTGGTGTCCGGTCATTGCTGATGAGAGAAATCAAAATGTGGGT 344
XX
QY 1161 CATATTCCGAAAGCCTTGGCCCTTGGGGCTCCACAGTCATGATGGCCTCTCTCGTGGCT 1220
DB 345 CATATTCCGAAAGCCTTGGCCCTTGGGGCTCCACAGTCATGATGGCCTCTCTCGTGGCT 404
XX
QY 1221 GCCACCACTGAGGCCCTCGGTGATGATCTCTTT 1253
DB 405 GCCACCACTGAGGCCCTCGGTGATGATCTCTTT 437
XX
RESULT 30
ABV05699
ID ABV05699 standard; cDNA: 389 BP.
XX
AC ABV05699;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 5690.
XX
KW Human; prostate cancer; cytosolic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
```

```
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 952; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 389 BP; 85 A; 97 C; 114 G; 92 T; 1 other:
XX
Query Match 6.8%; Score 112; DB 23; Length 389;
Best Local Similarity 100.0%; Pred. No. 3.7e-43;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 904 TGATCAAGTTCATCAAGAACAATACCTTAATCTCCAGTCATTGGAGGCAATGTGTCA 963
DB 28 TGATCAAGTTCATCAAGAACAATACCTTAATCTCCAGTCATTGGAGGCAATGTGTCA 87
XX
QY 964 CTGTGCGCCAGGCCAAGACCTCATTTGATGCGAGTGGCCCTCGGGGT 1015
DB 88 CTGTGCGCCAGGCCAAGACCTCATTTGATGCGAGTGGCCCTCGGGGT 139
XX
RESULT 31
AAI84483
ID AAI84483 standard; cDNA: 286 BP.
XX
AC AAI84483;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4543.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
```



```
XX DR WPI: 2001-514838/56.
XX P-PSDB: AA004552.
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders.
XX PS Claim 1: SEQ ID NO 4543; 1399pp + Sequence Listing: English.
XX CC The invention relates to human polynucleotides (AA17941-AA193841) and
XX CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoietic regulatory
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activity/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 286 BP: 78 A; 79 C; 60 G; 69 T; 0 other:

Query Match      4.2%; Score 69; DB 22; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY 278 GATGGCATATGCAATGGCTTACAGCGGTATGCTTATCACCACACTGTACACC 337
Db 157 GATGGCATATGCAATGGCTTACAGCGGTATGCTTATCACCACACTGTACACC 216
OY 338 TGAATTCCA 346
Db 217 TGAATTCCA 225

RESULT 32
AA003541
ID AA003541 standard; DNA; 1620 BP.
XX AC AA003541;
XX 31-AUG-1990 (first entry)
XX DE Chinese hamster IMPDH.
XX KM Inosine 5'-mono-phosphate dehydrogenase; hepatomas;
XX KM guanosine monophosphate; ss.
XX OS Cricetus sp.
XX PN WO9001545-A.
XX PD 22-FEB-1990.
XX PF 02-AUG-1989; 89MO-0000344.
XX PR 12-AUG-1988; 88US-0232302.
XX PA (ARCH-) ARCH. DEV. CORP.
XX PI Collart FR, Huberman E;
XX DR WPI: 1990-083504/11.
XX DR N-PSDB: AA003541.
XX PT DNA encoding eukaryotic inosine 5'-mono-phosphate dehydrogenase -
XX PT DNA encoding high levels of enzyme activity of tumor cells,
XX PT esp. hepatomas, and to produce guanosine monophosphate.
```

```
XX PS Claim 8; Fig 2: 51pp; English.
XX CC The sequence can be used to quantitatively detect IMPDH encoding
XX CC DNA or RNA by hybridisation, eg in normal and malignant cells.
XX CC esp. hepatomas.
XX CC See also AA003540.
XX SQ Sequence 1620 BP; 423 A; 369 C; 421 G; 406 T; 1 other:

Query Match      3.7%; Score 62; DB 11; Length 1620;
Best Local Similarity 100.0%; Pred. No. 2.8e-19;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY 306 GGTATTGGCTTCATCCACCACTGTACACCTGAATTCCAGGCCAATGAATTGGAAA 365
Db 284 GGTATTGGCTTCATCCACCACTGTACACCTGAATTCCAGGCCAATGAATTGGAAA 343
OY 366 GT 367
Db 344 GT 345

RESULT 33
ABN32859
ID ABN32859 standard; DNA; 60 BP.
XX AC ABN32859;
XX DT 15-JUL-2002 (first entry)
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:5607.
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX OS Homo sapiens.
XX PN WO200210449-A2.
XX PD 07-FEB-2002.
XX PF 20-JUL-2001; 2001WO-1B01903.
XX PR 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR WPI: 2002-257383/30.
XX PT New oligonucleotide libraries comprising oligonucleotides which
XX PT selectively hybridize to mRNAs transcribed from a transcription unit of
XX PT a genome, useful for detecting tissue-, pathology-, and
XX PT developmental-specific genes.
XX PS Example 1; SEQ ID 5607; 47pp; English.
XX CC The present invention describes oligonucleotide libraries for detecting
XX CC messenger RNAs that populate a (sub-)transcriptome, where the
XX CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX CC transcription units that populate a genome. The library comprises
XX CC several oligonucleotides, each capable of hybridising selectively to a
XX CC set of messenger RNAs transcribed from a given transcription unit of
XX CC the genome, which encodes one or more messenger RNA splice variants.
XX CC The oligonucleotide libraries are useful for detecting mRNAs from a
XX CC biological sample, in expression profiling studies, in qualitatively or
XX CC quantitatively characterising the corresponding transcriptome, and in
XX CC detecting RNA transcripts and splice variants of human or animal
XX CC transcriptomes. The libraries may also be used as specialised mini
XX CC libraries to detect transcripts of a sub-transcriptome under a
```

CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition: to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN59589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pubd/published\\_pcl\\_sequences](http://wipo.int/pubd/published_pcl_sequences).

XX Sequence 60 BP; 17 A; 17 C; 14 G; 12 T; 0 other;

Query Match 3.6%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred No. 2.6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 921 GACAAATACCCCTAATCTCAAGTCATTGAGGCAATGTGCTCACTGCTGCCAGGCCAAG 980  
DB 1 GACAAATACCCCTAATCTCAAGTCATTGAGGCAATGTGCTCACTGCTGCCAGGCCAAG 60

RESULT 34  
AAH90485/c  
ID AAH90485 standard; cDNA; 51 BP.

XX AAH90485;

DT 08-OCT-2001 (first entry)

XX Human clone c943922096 SNP site, SEQ ID NO:365.

XX Human: single nucleotide polymorphism: SNP: chromosome 3p24.2;  
KM detection; identification; gene therapy; genetic disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH replace(25..27,GC)  
FT /tag= a  
FT /standard\_name= "single nucleotide polymorphism"

XX WO200147942-A2.  
PN 05-JUL-2001.

XX PF 27-DEC-2000; 2000WO-US35387.  
PD 27-DEC-1999; 99US-0472865.

XX PR 27-DEC-1999; 99US-0472865.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinketsu RA, Leach M;

XX WPI; 2001-425617/45.

XX New polynucleotides containing single nucleotide polymorphisms, for  
PT detecting the presence of polymorphism, detecting a polymorphic site,  
PT and treating a patient suffering from a pathology ascribed to the  
PT polymorphism.

XX Claim 1; Page 99; 2955pp; English.

XX Sequences AAH90121-AAH90700 represent 580 human cDNA sequences which  
CC contain single nucleotide polymorphisms (SNPs). Sequences 1 to  
CC 568 (AAH90121-AAH90688) are consecutive pairs of nucleotides which  
CC contain silent SNPs. Sequences 569 to 580 (AAH90689-AAH90700) are  
CC consecutive pairs of nucleotides containing SNPs which result in changes  
CC in the corresponding amino acid sequences (AAG64751-AAG64762). The SNPs  
CC in sequences 569 to 574 (AAH90689-AAH90694) lead to conservative amino  
CC acid changes, while those in sequences 575 to 578 (AAH90695-AAH90698)  
CC result in non-conservative changes. The SNP in sequences 579 and 580

CC (AAH90699-AAH90700) generates a frameshift mutation. The invention also  
CC relates to a method of detecting a polymorphic site in a nucleic acid and  
CC a method of determining the relatedness of two nucleic acids. It also  
CC encompasses peptides containing polymorphic sites, antibodies raised  
CC against such peptides, and a method of detecting polymorphic proteins/  
CC peptides using the antibodies. The nucleic acids are useful for gene  
CC therapy of an individual having, suspected of having, or at risk of  
CC developing a pathological condition due to the presence of a sequence  
CC polymorphism. Such treatment would comprise administration of the  
CC wild-type nucleic acid sequence. Antibodies raised against polymorphic  
CC peptides can also be used in the treatment of such individuals.

XX Sequence 51 BP; 12 A; 20 C; 13 G; 6 T; 0 other;

Query Match 3.1%; Score 51; DB 22; Length 51;  
Best Local Similarity 100.0%; Pred No. 5.2e-14;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CGCGCGTGTCTGTGTGGCCATGCGCACTACCTGATTAGTGGGGACAG 77  
DB 51 CGCGCGTGTCTGTGTGGCCATGCGCACTACCTGATTAGTGGGGACAG 1

RESULT 35  
AAT25471  
ID AAT25471 standard; cDNA to mRNA; 50 BP.

XX AAT25471;

DT 17-OCT-1996 (first entry)

XX Human gene signature HUMGS07638.

XX Gene signature: messenger RNA; mRNA; relative abundance; frequency;  
KM human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX MO9514772-A1.  
PN 01-JUN-1995.

XX PF 11-NOV-1994; 94WO-JP01916.  
PD 11-NOV-1993; 93JP-0355504.

XX PR 12-NOV-1993; 93JP-0355504.

XX PA (MATS/) MATSUBARA K.  
XX (OKUBO/) OKUBO K.

XX PI Matsubara K, Okubo K;

XX WPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues

XX Claim 1; Page 1838; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7897 "GS" sequences  
CC given in AAT19001-T26837 and which is able to hybridize to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.  
XX  
SQ Sequence 50 BP: 17 A; 10 C; 7 G; 16 T; 0 other;  
Query Match 2.9%; Score 48; DB 16; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1,4e-12;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1598 GATTCAGCACACCTCTCGGCTTTTTCATTAAGATTAGAAAG 1645  
1 GATTCAGCACACCTCTCGGCTTTTTCATTAAGATTAGAAAG 48  
Db  
RESULT 36  
ABK40085  
ID ABK40085 standard; DNA: 6193 BP.  
XX  
AC ABK40085;  
XX  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE Human chemically pretreated gene sequence #84 strand 1.  
XX  
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;  
KW cytostatic; ALDH6; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;  
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
PN MO200202806-A2.  
XX  
PD 10-JAN-2002.  
XX  
PF 29-JUN-2001: 2001WO-EP07470.  
XX  
PR 30-JUN-2000: 2000DE-1032529.  
PR 01-SEP-2000: 2000DE-1043826.  
XX  
PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI: 2002-154757/20.  
XX  
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,  
PT useful for detecting cytosine methylation state of genes associated  
PT with pharmacogenomics and for therapy of diseases e.g. cancer -  
XX  
PS Claim 1: SEQ ID No 167; 24pp; English.  
XX  
CC The invention relates to a nucleic acid comprising a sequence at  
CC least 18 bases in length of a segment of the chemically pretreated DNA  
CC of genes associated with pharmacogenomics according to one of the  
CC sequences of the genes ALDH6 (NM\_000693), CYP11A (NM\_000781), CYP11B1  
CC (NM\_000497), CYP3A3 (NM\_000776 and NM\_017460), DPYD (NM\_000110), EPHX2  
CC (NM\_001979), OCLN (NM\_002538), TXNRD1 (NM\_003330), UGT8 (NM\_003360),  
CC MRP (NM\_004996, NM\_019900, NM\_019901, NM\_019902, NM\_019862, NM\_019898,  
CC NM\_019899) and their complementary sequences, or a sequence (SI) chosen  
CC from 87 sequences and their complements. The chemical pretreatment  
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)  
CC into uracils. Also included are an oligomer (II) in particular an  
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in  
CC each case at least one base sequence having a length of 9 nucleotides  
CC which hybridises to or is identical to a chemically pretreated DNA of  
CC genes associated with pharmacogenomics and their complements, arranged in  
CC an array for analysing diseases associated with the methylation state  
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)  
CC of the 87 sequences. The oligomers may also be used as PCR primers.  
CC The set of 87 nucleic acids and their complements is useful for diagnosis  
CC and therapy of solid tumours and cancer. The present sequence  
CC represents one the 87 DNA sequences or its complement

CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 6193 BP: 1404 A; 124 C; 1628 G; 3037 T; 0 other;  
Query Match 1.7%; Score 28; DB 24; Length 6193;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 417 AAGGATCGCGTGGGATGTTTGAAG 444  
1 AAGGATCGCGTGGGATGTTTGAAG 2588  
Db 2561 AAGGATCGCGTGGGATGTTTGAAG 2588  
RESULT 37  
AAH90486/c  
ID AAH90486 standard; cDNA: 50 BP.  
XX  
XX  
AC AAH90486;  
XX  
DT 08-OCT-2001 (first entry)  
XX  
DE Human clone c943922096 SNP site, SEQ ID NO:366.  
XX  
KW Human; single nucleotide polymorphism; SNP; chromosome 3p24.2;  
KW detection; identification; gene therapy; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key location/Qualifiers  
FH variation replace(25..26,GGC)  
FT /\*tag= a  
FT /standard\_name= "single nucleotide polymorphism"  
XX  
PN MO200147942-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 27-DEC-2000: 2000WO-US35387.  
XX  
PR 27-DEC-1999: 99US-0472865.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI: 2001-425617/45.  
XX  
PT New polynucleotides containing single nucleotide polymorphisms, for  
PT detecting the presence of polymorphism, detecting a polymorphic site,  
PT and treating a patient suffering from a pathology ascribed to the  
PT polymorphism -  
XX  
PS Claim 1: Page 95; 295pp; English.  
XX  
CC Sequences AAH90121-AAH90700 represent 580 human cDNA sequences which  
CC contain single nucleotide polymorphisms (SNPs). Sequences 1 to  
CC 568 (AAH90121-AAH90688) are consecutive pairs of nucleotides which  
CC contain silent SNPs. Sequences 569 to 580 (AAH90689-AAH90700) are  
CC consecutive pairs of nucleotides containing SNPs which result in changes  
CC in the corresponding amino acid sequences (AAG64751-AAG64762). The SNPs  
CC in sequences 569 to 574 (AAH90689-AAH90694) lead to conservative amino  
CC acid changes, while those in sequences 575 to 578 (AAH90695-AAH90698)  
CC result in non-conservative changes. The SNP in sequences 579 and 580  
CC (AAH90699-AAH90700) generates a frameshift mutation. The invention also  
CC relates to a method of detecting a polymorphic site in a nucleic acid and  
CC a method of determining the relatedness of two nucleic acids. It also  
CC encompasses peptides containing polymorphic sites, antibodies raised  
CC against such peptides, and a method of detecting polymorphic proteins/  
CC peptides using the antibodies. The nucleic acids are useful for gene  
CC therapy of an individual having, suspected of having, or at risk of

CC developing a pathological condition due to the presence of a sequence  
 CC polymorphism. Such treatment would comprise administration of the  
 CC wild-type nucleic acid sequence. Antibodies raised against polymorphic  
 CC peptides can also be used in the treatment of such individuals.  
 XX

SQ Sequence 50 BP; 12 A; 20 C; 12 G; 6 T; 0 other;

Query Match 1.6%; Score 26; DB 22; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 0.046;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CGCGGTCCTCTGTTGGCCATGCG 52  
 |||||||  
 DB 50 CGCGGTCCTCTGTTGGCCATGCG 25

RESULT 38  
 AAH90484/c  
 ID AAH90484 standard; cDNA; 51 BP.

XX AAH90484;

DT 08-OCT-2001 (first entry)

XX Human clone c943922096 SNP site, SEQ ID NO:364.

XX Human: single nucleotide polymorphism; SNP; chromosome 3p24.2;  
 KM detection; identification; gene therapy; genetic disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT variation replace(26,A)  
 FT /tag= a

FT /standard\_name= "single nucleotide polymorphism"

PN WO200147942-A2.

PD 05-JUL-2001.

PF 27-DEC-2000: 2000WO-US5387.

PR 27-DEC-1999: 990S-0472865.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI: 2001-425617/45.

PT New polynucleotides containing single nucleotide polymorphisms, for  
 PT detecting the presence of polymorphism, detecting a polymorphic site,  
 PT and treating a patient suffering from a pathology ascribed to the  
 PT polymorphism -

PS Claim 1; Page 99; 295bp; English.

XX Sequences AAH90121-AAH90700 represent 580 human cDNA sequences which  
 CC contain single nucleotide polymorphisms (SNPs). Sequences 1 to  
 CC 568 (AAH90121-AAH90688) are consecutive pairs of nucleotides which  
 CC contain silent SNPs. Sequences 569 to 580 (AAH90689-AAH90700) are  
 CC consecutive pairs of nucleotides containing SNPs which result in changes  
 CC in the corresponding amino acid sequences (AAG64751-AAG64762). The SNPs  
 CC in sequences 569 to 574 (AAH90689-AAH90694) lead to conservative amino  
 CC acid changes, while those in sequences 575 to 578 (AAH90695-AAH90698)  
 CC result in non-conservative changes. The SNP in sequences 579 and 580  
 CC (AAH90699-AAH90700) generates a frameshift mutation. The invention also  
 CC relates to a method of detecting a polymorphic site in a nucleic acid and  
 CC a method of determining the relatedness of two nucleic acids. It also  
 CC encompasses peptides containing polymorphic sites, antibodies raised  
 CC against such peptides, and a method of detecting polymorphic proteins/  
 CC peptides using the antibodies. The nucleic acids are useful for gene  
 CC therapy of an individual having, suspected of having, or at risk of

CC developing a pathological condition due to the presence of a sequence  
 CC polymorphism. Such treatment would comprise administration of the  
 CC wild-type nucleic acid sequence. Antibodies raised against polymorphic  
 CC peptides can also be used in the treatment of such individuals.  
 XX

SQ Sequence 51 BP; 6 A; 18 C; 12 G; 15 T; 0 other;

Query Match 1.6%; Score 26; DB 22; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 0.046;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 AGAGGGAAGACTTGTTGATAGCCCC 614  
 |||||||  
 DB 51 AGAGGGAAGACTTGTTGATAGCCCC 26

RESULT 39  
 AAD28933/c  
 ID AAD28933 standard; DNA; 35 BP.

XX AAD28933;

DT 07-MAY-2002 (first entry)

XX Human wild-type, type II IMPDH cDNA amplifying 3' primer.

XX Human: inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;  
 KM proliferative-type disease; cancer; imaging methodology; cytostatic;  
 KW therapy; PCR primer; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT variation replace(26,A)  
 FT /tag= a

PN WO200185952-A2.

PD 15-NOV-2001.

PF 10-MAY-2001: 2001WO-US15457.

PR 10-MAY-2000: 2000US-203448P.

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;

PI Mouravieff JE, Einspahr HM, Kish K;

DR WPI: 2002-164105/21.

PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
 PT having an oligo-peptide domain substituted for a subdomain of a  
 PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
 PT antibodies -

PS Example 1; Page 56; 161pp; English.

XX The invention relates to modified inosine 5'-monophosphate dehydrogenase  
 CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
 CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
 CC modified IMPDH polypeptides are useful for drug discovery for  
 CC therapeutic, diagnostic and prognostic procedures for detecting or  
 CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
 CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
 CC diagnostic and prognostic markers of diseases, as targets for various  
 CC therapeutic modalities, and to identify and isolate ligands and other  
 CC agents that bind to modified IMPDH. These antibodies may be used in  
 CC diagnostic assays, imaging methodologies, therapeutic methods in the  
 CC management of cancer or other proliferative-type diseases, and in  
 CC purifying modified IMPDH polypeptides and for isolating related  
 CC molecules such as wild type and mutant IMPDH polypeptides. The present  
 CC sequence is a PCR primer used to amplify human wild-type, type II  
 CC IMPDH cDNA.

SQ Sequence 35 BP; 10 A; 9 C; 7 G; 9 T; 0 other;

Query Match 1.5%; Score 25; DB 24; Length 35;  
Best Local Similarity 100.0%; Pred.No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1570 CGTATGAGAACGGCTTTCTGTGAAA 1594  
|||||  
Db 35 CGTATGAGAACGGCTTTCTGTGAAA 11

```
RESULT 40
AAH90483/c
ID AAH90483 standard; cDNA: 51 BP.
XX
AC AAH90483;
XX
DT 08-OCT-2001 (first entry)
XX
DE Human clone cg43922096 SNP site, SEQ ID NO:363.
XX
KW Human: single nucleotide polymorphism; SNP: chromosome 3p24.2;
KW detection: identification: gene therapy; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(26,G)
FT /*tag= a
/standard_name= "single nucleotide polymorphism"
PN WO200147942-A2.
XX
PD 05-JUL-2001.
XX
PF 27-DEC-2000: 2000WO-US35387.
XX
PR 27-DEC-1999: 99US-0472865.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI: 2001-425617/45.
XX
PT New polynucleotides containing single nucleotide polymorphisms, for
PT detecting the presence of polymorphism, detecting a polymorphic site,
PT and treating a patient suffering from a pathology ascribed to the
PT polymorphism
XX
PS Claim 1: Page 99; 295pp; English.
XX
CC Sequences AAH90121-AAH90700 represent 580 human cDNA sequences which
CC contain single nucleotide polymorphisms (SNPs). Sequences 1 to
CC 568 (AAH90121-AAH90688) are consecutive pairs of nucleotides which
CC contain silent SNPs. Sequences 569 to 580 (AAH90689-AAH90700) are
CC consecutive pairs of nucleotides containing SNPs which result in changes
CC in the corresponding amino acid sequences (AAG64751-AG64762). The SNPs
CC in sequences 569 to 574 (AAH90689-AAH90694) lead to conservative amino
CC acid changes, while those in sequences 575 to 578 (AAH90695-AAH90698)
CC result in non-conservative changes. The SNP in sequences 579 and 580
CC (AAH90699-AAH90700) generates a frameshift mutation. The invention also
CC relates to a method of detecting a polymorphic site in a nucleic acid and
CC a method of determining the relatedness of two nucleic acids. It also
CC encompasses peptides containing polymorphic sites, antibodies raised
CC against such peptides, and a method of detecting polymorphic proteins/
CC peptides using the antibodies. The nucleic acids are useful for gene
CC therapy of an individual having, suspected of having, or at risk of
CC developing a pathological condition due to the presence of a sequence
CC polymorphism. Such treatment would comprise administration of the
CC wild-type nucleic acid sequence. Antibodies raised against polymorphic
CC peptides can also be used in the treatment of such individuals.
XX
SQ Sequence 51 BP: 7 A: 18 C: 11 G: 15 T: 0 other:
```

```
Query Match 1.5%; Score 25; DB 22; Length 51;
Best Local Similarity 100.0%; Pred.No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 589 AGAGGAGACACTGTGTGTAGCCCC 613
|||||
Db 51 AGAGGAGACACTGTGTGTAGCCCC 27

RESULT 41
AAA30019/c
ID AAA30019 standard; DNA: 27 BP.
XX
AC AAA30019;
XX
DT 09-AUG-2000 (first entry)
XX
DE Hairpin hybridizer molecule #24 targeting c-rat and IMPDH RNA.
XX
KW DNA-RNA hybrid: phosphorothioate; hairpin hybridizer molecule; HPH;
KW c-rat; inhibitor; detect; nuclease resistance; gene function; IMPDH;
KW gene expression modulator; inosine 5'-monophosphate dehydrogenase; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_RNA 1..7
FT /*tag= a
/note= "2-O-methyl ribonucleotides"
FT stem_loop 1..27
FT /*tag= b
FT misc_feature 8..17
FT /*tag= c
/note= "Phosphorothioate internucleotide linkage"
FT misc_RNA 17..27
FT /*tag= d
/note= "2-O-methyl ribonucleotides"
PN WO200017346-A2.
XX
PD 30-MAR-2000.
XX
PF 20-SEP-1999: 99WO-US21865.
XX
PR 21-SEP-1998: 98US-0101174.
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
PI Hartmann T, Zwick M, Thompson J, Jarvis T;
XX
DR WPI: 2000-292841/25.
XX
PT Modulating target sequence in a cell, useful e.g. therapeutically or
PT for identifying gene function, by treatment with novel hairpin
PT hybridizer nucleic acid molecules
XX
PS Claim 101; Page 77; 128pp; English.
XX
CC A method for modulating the function of a target sequence in a cell
CC comprises treatment with a hairpin hybridizer (HPH) nucleic acid molecule
CC such as that represented by the present sequence. The present HPH
CC molecule is used in an example of the invention and targets the human
CC c-rat and inosine 5'-monophosphate dehydrogenase (IMPDH) RNA. The HPH
CC molecules function through RNA/DNA inhibition. The HPH molecule binds to
CC and blocks the function of a target nucleic acid, and modulated cellular
CC and viral processes such as splicing, editing, and translation. The HPH
CC molecule can be used therapeutically, in target validation, to identify
CC gene function and/or therapeutic targets, for analysis of mutations in
CC resistance cell and to detect specific RNA. The hairpin structure improves
CC resistance to nuclease degradation, localization within the cell, and
CC uptake by cells. The HPH may include a sequence that activates RNase H
CC (for cleaving RNA) and its specificity is greater than for linear
CC antisense sequences.
```

XX Sequence 27 BP: 7 A; 9 C; 5 G; 3 T; 3 U; 0 other;  
SQ  
Query Match 1.5%; Score 24; DB 21; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 764 TGCCAGAGAAACAGCTGCTGTGTGG 787  
|||||  
24 TGCCAGAGAAACAGCTGCTGTGTGG 1  
Db  
RESULT 42  
AAA30020/c  
ID AAA30020 standard; DNA; 29 BP.  
XX  
AC AAA30020;  
XX  
DT 09-AUG-2000 (first entry)  
XX  
DE Hairpin hybridizer molecule #25 targeting c-raf and IMPDH RNA.  
XX  
KM DNA-RNA hybrid; phosphorothioate; hairpin hybridizer molecule; HPH;  
KM c-raf; inhibitor; detect; nuclease resistance; gene function; IMPDH;  
KM gene expression modulator; inosine 5'-monophosphate dehydrogenase; ss.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT misc-RNA 1..7  
FT /\*tag= a  
FT /\*note= "2-O-methyl ribonucleotides"  
FT stem\_loop 1..29  
FT /\*tag= b  
FT misc-feature 8..17  
FT /\*tag= c  
FT /\*note= "Phosphorothioate internucleotide linkage"  
FT misc-RNA 17..29  
FT /\*tag= d  
FT /\*note= "2-O-methyl ribonucleotides"  
XX  
PN WO200017346-A2.  
XX  
PD 30-MAR-2000.  
XX  
PE 20-SEP-1999; 99WO-US21865.  
XX  
PR 21-SEP-1998; 98US-0101174.  
XX  
PA (RIBO-) RIBOZYME PHARM INC.  
XX  
PI Hartmann T, Zwick M, Thompson J, Jarvis T;  
XX  
DR WPI: 2000-292841/25.  
XX  
PT Modulating target sequence in a cell, useful e.g. therapeutically or  
PT for identifying gene function, by treatment with novel hairpin  
PT hybridizer nucleic acid molecules -  
XX  
PS Claim 101; Page 77; 128pp; English.  
XX  
CC A method for modulating the function of a target sequence in a cell  
CC comprises treatment with a hairpin hybridizer (HPH) nucleic acid molecule  
CC such as that represented by the present sequence. The present HPH  
CC molecule is used in an example of the invention and targets the human  
CC c-raf and inosine 5'-monophosphate dehydrogenase (IMPDH) RNA. The HPH  
CC molecules function through RNA/DNA inhibition. The HPH molecule binds to  
CC and blocks the function of a target nucleic acid, and modulated cellular  
CC and vital processes such as splicing, editing, and translation. The HPH  
CC molecule can be used therapeutically, in target validation, to identify  
CC gene function and/or therapeutic targets, for analysis of mutations in  
CC diseased cell and to detect specific RNA. The hairpin structure improves  
CC resistance to nuclease degradation, localization within the cell, and

CC uptake by cells. The HPH may include a sequence that activates RNase H  
CC (for cleaving RNA) and its specificity is greater than for linear  
CC antisense sequences.  
XX  
SQ Sequence 29 BP: 7 A; 10 C; 6 G; 3 T; 3 U; 0 other;  
SQ  
Query Match 1.5%; Score 24; DB 21; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 764 TGCCAGAGAAACAGCTGCTGTGTGG 787  
|||||  
24 TGCCAGAGAAACAGCTGCTGTGTGG 1  
Db  
RESULT 43  
AAD28932  
ID AAD28932 standard; DNA; 33 BP.  
XX  
AC AAD28932;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Human wild-type, type II IMPDH cDNA amplifying 5' primer.  
XX  
KM Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;  
KM proliferative-type disease; cancer; imaging methodology; cytostatic;  
KM therapy; PCR primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200185952-A2.  
XX  
PD 15-NOV-2001.  
XX  
PE 10-MAY-2001; 2001WO-US15457.  
XX  
PR 10-MAY-2000; 2000US-203448P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff JE, Einspahr HM, Kish K;  
XX  
DR WPI: 2002-164105/21.  
XX  
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
PS Example 1; Page 56; 161pp; English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is a PCR primer used to amplify human wild-type, type II  
CC IMPDH cDNA.  
XX  
SQ Sequence 33 BP: 7 A; 8 C; 9 G; 9 T; 0 other;  
SQ  
Query Match 1.5%; Score 24; DB 24; Length 33;

Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 24: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 48 ATGCCGACTGCTGATTAGTGGG 71  
|  
10 ATGCCGACTGCTGATTAGTGGG 33

RESULT 44  
AAA30021/C  
ID AAA30021 standard: DNA: 23 BP.

AC AAA30021:

DT 09-AUG-2000 (first entry)

XX Hairpin hybridizer molecule #24 targeting c-raf and IMPDH RNA.

DE DNA-RNA hybrid: phosphorothioate: hairpin hybridizer molecule: HPH;  
KW c-raf; inhibitor; detect; nuclease resistance; gene function: IMPDH;  
KW gene expression modulator: inosine 5-monophosphate dehydrogenase; ss.

XX Synthetic.

FT Key Location/Qualifiers  
FT misc\_RNA 1..7

FT /\*tag= a /note= "2-O-methyl ribonucleotides"

FT stem\_loop 1..23 /\*tag= b

FT misc\_feature 8..17 /\*tag= c

FT /\*note= "Phosphorothioate internucleotide linkage"

FT /\*tag= d /note= "2-O-methyl ribonucleotides"

PN WO200017346-A2.

XX 30-MAR-2000.

XX 20-SEP-1999: 99WO-US21865.

XX 21-SEP-1998: 98US-0101174.

XX (RIBO-) RIBOZYME PHARM INC.

XX Hartmann T, Zwick M, Thompson J, Jarvis T;

XX WPI: 2000-292841/25.

XX Modulating target sequence in a cell, useful e.g. therapeutically or  
XX for identifying gene function, by treatment with novel hairpin  
XX hybridizer nucleic acid molecules

XX Example 6; Page 77; 128pp: English.

XX A method for modulating the function of a target sequence in a cell  
XX comprises treatment with a hairpin hybridizer (HPH) nucleic acid molecule  
XX such as that represented by the present sequence. The present HPH  
XX molecule is used in an example of the invention and targets the human  
XX c-raf and inosine 5-monophosphate dehydrogenase (IMPDH) RNA. The HPH  
XX molecules function through RNA/DNA inhibition. The HPH molecule binds to  
XX and blocks the function of a target nucleic acid, and modulated cellular  
XX and viral processes such as splicing, editing, and translation. The HPH  
XX molecule can be used therapeutically, in target validation, to identify  
XX gene function and/or therapeutic targets, for analysis of mutations in  
XX diseased cell and to detect specific RNA. The hairpin structure improves  
XX resistance to nuclease degradation, localization within the cell, and  
XX uptake by cells. The HPH may include a sequence that activates RNase H  
XX (for cleaving RNA) and its specificity is greater than for linear  
XX antisense sequences.

SO Sequence 23 BP: 1 A; 6 C; 9 G; 3 T; 4 U; 0 other;

Query Match 1.48; Score 23; DB 21; Length 23;

Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 23: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 CGACGACTCAGACAGCAGC 114  
|  
23 CGACGACTCAGACAGCAGC 1

RESULT 45  
AAA30022/C  
ID AAA30022 standard: DNA: 27 BP.

AC AAA30022:

DT 09-AUG-2000 (first entry)

XX Hairpin hybridizer molecule #27 targeting c-raf and IMPDH RNA.

DE DNA-RNA hybrid: phosphorothioate: hairpin hybridizer molecule: HPH;  
KW c-raf; inhibitor; detect; nuclease resistance; gene function: IMPDH;  
KW gene expression modulator: inosine 5-monophosphate dehydrogenase; ss.

XX Synthetic.

FT Key Location/Qualifiers  
FT misc\_RNA 1..7

FT /\*tag= a /note= "2-O-methyl ribonucleotides"

FT stem\_loop 1..27 /\*tag= b

FT misc\_feature 8..17 /\*tag= c

FT /\*note= "Phosphorothioate internucleotide linkage"

FT /\*tag= d /note= "2-O-methyl ribonucleotides"

PN WO200017346-A2.

XX 30-MAR-2000.

XX 20-SEP-1999: 99WO-US21865.

XX 21-SEP-1998: 98US-0101174.

XX (RIBO-) RIBOZYME PHARM INC.

XX Hartmann T, Zwick M, Thompson J, Jarvis T;

XX WPI: 2000-292841/25.

XX Modulating target sequence in a cell, useful e.g. therapeutically or  
XX for identifying gene function, by treatment with novel hairpin  
XX hybridizer nucleic acid molecules

XX Claim 101; Page 77; 128pp: English.

XX A method for modulating the function of a target sequence in a cell  
XX comprises treatment with a hairpin hybridizer (HPH) nucleic acid molecule  
XX such as that represented by the present sequence. The present HPH  
XX molecule is used in an example of the invention and targets the human  
XX c-raf and inosine 5-monophosphate dehydrogenase (IMPDH) RNA. The HPH  
XX molecules function through RNA/DNA inhibition. The HPH molecule binds to  
XX and blocks the function of a target nucleic acid, and modulated cellular  
XX and viral processes such as splicing, editing, and translation. The HPH  
XX molecule can be used therapeutically, in target validation, to identify  
XX gene function and/or therapeutic targets, for analysis of mutations in  
XX diseased cell and to detect specific RNA. The hairpin structure improves  
XX resistance to nuclease degradation, localization within the cell, and  
XX uptake by cells. The HPH may include a sequence that activates RNase H







Query Match 1.3%; Score 22; DB 24; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: February 12, 2003, 23:43:00  
Job time : 374 secs

OY 357 GTTCGGAAGTGAAGAAATATG 378  
|||||  
DB 30 GTTCGGAAGTGAAGAAATATG 9

RESULT 50  
AAD28939/C  
ID AAD28939 standard; DNA: 33 BP.  
XX  
AC AAD28939;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Human type II AGRP-IMPDH cDNA amplifying 3' primer.  
XX  
KM Human: inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;  
KM proliferative-type disease; cancer; imaging methodology; cytostatic;  
KM therapy; PCR primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200185952-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 10-MAY-2001; 2001WO-US15457.  
XX  
PR 10-MAY-2000; 2000US-203448P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PI Krystek SR, Sherif S, Wilmer MR, Hollenbaugh DL, Yan N;  
PI Mouravieff JE, Einspahr HM, Kish K;  
XX  
DR WPI; 2002-164105/21.  
XX  
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide  
PT having an oligo-peptide domain substituted for a subdomain of a  
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating  
PT antibodies -  
XX  
PS Example 1; Page 58; 161pp; English.  
XX  
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase  
CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain  
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The  
CC modified IMPDH polypeptides are useful for drug discovery, for  
CC therapeutic, diagnostic and prognostic procedures for detecting or  
CC quantifying modified IMPDH polypeptides and their corresponding nucleic  
CC acids. IMPDH polypeptides are also useful for generating antibodies, as  
CC diagnostic and prognostic markers of diseases, as targets for various  
CC therapeutic modalities, and to identify and isolate ligands and other  
CC agents that bind to modified IMPDH. These antibodies may be used in  
CC diagnostic assays, imaging methodologies, therapeutic methods in the  
CC management of cancer or other proliferative-type diseases, and in  
CC purifying modified IMPDH polypeptides and for isolating related  
CC molecules such as wild type and mutant IMPDH polypeptides. The present  
CC sequence is a PCR primer used to amplify human type II AGRP-IMPDH cDNA.  
XX  
SQ Sequence 33 BP; 8 A; 11 C; 5 G; 9 T; 0 other;

Query Match 1.3%; Score 22; DB 24; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 357 GTTCGGAAGTGAAGAAATATG 378  
|||||  
DB 33 GTTCGGAAGTGAAGAAATATG 12

[illegible]

PF 02-AUG-1989; 89WO-0000344.  
XX  
PR 12-AUG-1988; 88US-0232302.  
XX  
PA (ARCH-) ARCH. DEV. CORP.  
XX  
PI Collart FR, Huberman E;  
XX  
DR WPI; 1990-083504/11.  
P-PSDB: AAR05432.  
XX  
PT DNA encoding eukaryotic inosine 5'-mono-phosphate dehydrogenase -  
PT used to detect high levels of enzyme activity of tumour cells,  
PT esp. hepatomas, and to produce guanosine monophosphate.  
XX  
PS Claim 8; Fig 1; 51pp; English.  
XX  
CC The sequence can be used to quantitatively detect IMPDH encoding  
CC DNA or RNA by hybridisation, eg in normal and malignant cells,  
CC esp. hepatomas.  
CC See also AA003541.  
XX  
SQ Sequence 1642 BP; 411 A; 416 C; 449 G; 366 T; 0 other;

Query Match 98.9%; Score 1635.6; DB 11; Length 1642;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 GGGCGGTCTCTGGAGACAGCGCGGTGTCCTGTGTGGCCATGGCCGACTACTGATTA 66  
DB 1 GGGCGGTCTCTGGAGACAGCGCGGTGTCCTGTGTGGCCATGGCCGACTACTGATTA 60  
QY 67 GTGGGGGCACTGTCCTACGTCCAGACAGAGGACTCAGACAGAGCTCTTCACTGCG 126  
DB 61 GTGGGGGCACTGTCCTACGTCCAGAGAGAGGACTCAGACAGAGCTCTTCACTGCG 120  
QY 127 GAGAGCGCTCAGCTCAATGACTTTCATTCCTCCCTGGGTACATCGACTTCGACG 186  
DB 121 GAGAGCGCTCAGCTCAATGACTTTCATTCCTCCCTGGGTACATCGACTTCGACG 180  
QY 187 ACCAGGTGACGCTGACTTGTGCTGACCAAGAAATCACTTTAGACCCCACTGGTT 246  
DB 181 ACCAGGTGACGCTGACTTGTGCTGACCAAGAAATCACTTTAGACCCCACTGGTT 240  
QY 247 CCTCTCCATGAGACAGACAGTACAGAGGCTGGATGGCCATAGCAATGGCGTTACAGCG 306  
DB 241 CCTCTCCATGAGACAGTACAGAGGCTGGATGGCCATAGCAATGGCGTTACAGCG 300  
QY 307 GTATTGGCTTCATCCACCACTGTACACTGTAATTCAGGGCAATGAAGTTGGAAG 366  
DB 301 GTATTGGCTTCATCCACCACTGTACACTGTAATTCAGGGCAATGAAGTTGGAAG 360  
QY 367 TGAAGAAATATGACAGGGGATTTCATCAGACCCCTGGTCCAGCCCAAGGATCGCG 426  
DB 361 TGAAGAAATATGACAGGGGATTTCATCAGACCCCTGGTCCAGCCCAAGGATCGCG 420  
QY 427 TCGGGGATTTTTTGAAGCAAGGCCCGCATGTTTCTGCGGTATCCCAATCACAGACA 486  
DB 421 TCGGGGATTTTTTGAAGCAAGGCCCGCATGTTTCTGCGGTATCCCAATCACAGACA 480  
QY 487 CAGGCCGATGGGAGCGCGCTGGTGGGCATCTCTCCAGGAGCATGTGATTTTCTCA 546  
DB 481 CAGGCCGATGGGAGCGCGCTGGTGGGCATCTCTCCAGGAGCATGTGATTTTCTCA 540  
QY 547 AAGAGAGGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGAAAGATTGTGTG 606  
DB 541 AAGAGAGGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGAAAGATTGTGTG 600  
QY 607 TAGCCCCCGCAGCATCACTGTAAGAGAGCAAAATGAAATTTCTGAGCCGACAGAAAG 666  
DB 601 TAGCCCCCGCAGCATCACTGTAAGAGAGCAAAATGAAATTTCTGAGCCGACAGAAAG 660  
QY 667 GAAAGTTGCCCATTTGAAATGAAGATGATGAGCTTGGCCATCATTTGCCCGACAGACC 726

DB 661 GAAAGTTGCCCATTTGAAATGAAGATGATGAGCTTGTGCCATCTGCCGACAGACC 720  
QY 727 TGAAGAGAAATCGGGACATACCCACATAGCCTCCAAAGATGCCAAGAAACAGCGCTGTG 786  
DB 721 TGAAGAGAAATCGGGACATACCCACATAGCCTCCAAAGATGCCAAGAAACAGCGCTGTG 780  
QY 787 GGGCAGCCATTTGGCACTCATGAGATGACAACTATAGCTGACTTGTGCGCCAGGCTG 846  
DB 781 GGGCAGCCATTTGGCACTCATGAGATGACAACTATAGCTGACTTGTGCGCCAGGCTG 840  
QY 847 GTGTGATGTAGTGGTTTTTGGACTCTTCCAGGGAATTCATCTTCCAGATCAATATGA 906  
DB 841 GTGTGATGTAGTGGTTTTTGGACTCTTCCAGGGAATTCATCTTCCAGATCAATATGA 900  
QY 907 TCAAGTACATCAAGACAAATACCCCAATCTCCAGTCACTTGGAGGCAATGTGTCACGT 966  
DB 901 TCAAGTACATCAAGACAAATACCCCAATCTCCAGTCACTTGGAGGCAATGTGTCACGT 960  
QY 967 CTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGGATGCCCTGGCGGTGGCATGGGA 1026  
DB 961 CTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGGATGCCCTGGCGGTGGCATGGGA 1020  
QY 1027 GTGGCTCCATCTGCAATTCACAGAAATGCTGGCCTGTGGCGGCCCAAGCACAGAC 1086  
DB 1021 GTGGCTCCATCTGCAATTCACAGAAATGCTGGCCTGTGGCGGCCCAAGCACAGAC 1080  
QY 1087 TGTACAGGTGTATGATGTGACAGCGCGTTGGTTCGGTCAATGTGTATGAGGAA 1146  
DB 1081 TGTACAGGTGTATGATGTGACAGCGCGTTGGTTCGGTCAATGTGTATGAGGAA 1140  
QY 1147 TCCAAATGTGGGTATATATGCGAAAGCGCTTGGCCCTTCACAGTCAATGATGG 1206  
DB 1141 TCCAAATGTGGGTATATATGCGAAAGCGCTTGGCCCTTCACAGTCAATGATGG 1200  
QY 1207 GCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATCTTCTTTTCCATGGGATCC 1266  
DB 1201 GCTCTCTCTGCTGCCACACTGAGGCCCTGTGTGAATCTTCTTTTCCATGGGATCC 1260  
QY 1267 GGTAAAGAAATATCGCGGTATGGGTTCTCTGATGCCATGAGCAAGCACTCAGACGCC 1326  
DB 1261 GGTAAAGAAATATCGCGGTATGGGTTCTCTGATGCCATGAGCAAGCACTCAGACGCC 1320  
QY 1327 AGAAGATATTTCAAGTGAAGTGCACAAATCAAAAGTGGCCAGGAGTGTGTGCTG 1386  
DB 1321 AGAAGATATTTCAAGTGAAGTGCACAAATCAAAAGTGGCCAGGAGTGTGTGCTG 1380  
QY 1387 TGCAGACAAAGGTCATCCACAAATTTGTCCCTTACTGATTTGCTGGCATCCAACT 1446  
DB 1381 TGCAGACAAAGGTCATCCACAAATTTGTCCCTTACTGATTTGCTGGCATCCAACT 1440  
QY 1447 CATGCCAGACATTTGGTGCACAAGAGCTTGACCAAGTCCGAGCCATGATCTCTGGGG 1506  
DB 1441 CATGCCAGACATTTGGTGCACAAGAGCTTGACCAAGTCCGAGCCATGATCTCTGGGG 1500  
QY 1507 AGCTTAAGTTTGAAGAGAACAGCTCTCAGCCAGGTGGAAGTGGCGTCCATAGCCTCC 1566  
DB 1501 AGCTTAAGTTTGAAGAGAACAGCTCTCAGCCAGGTGGAAGTGGCGTCCATAGCCTCC 1560  
QY 1567 ATTTCGATGAGAAAGCGGCTTTTCTGAAAAGGATCCAGCACCTCTCGGTTTTTTTT 1626  
DB 1561 ATTTCGATGAGAAAGCGGCTTTTCTGAAAAGGATCCAGCACCTCTCGGTTTTTTTT 1620  
QY 1627 CAATAAAGTTTGAAGAACGCC 1648  
DB 1621 CAATAAAGTTTGAAGAACGCC 1642

RESULT 2  
ABV29489  
ID ABV29489 standard: cDNA; 2277 BP.  
XX  
AC ABV29489;

XX 16-SEP-2002 (first entry)  
XX  
XX Human prostate expression marker cDNA 29480.  
DE  
XX Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
PN MO200160860-A2.  
XX  
XX 23-AUG-2001.  
PD  
XX 20-FEB-2001: 2001WO-US05171.  
PF  
XX 17-FEB-2000: 2000US-183319P.  
PR 16-MAR-2000: 2000US-189862P.  
XX 25-MAY-2000: 2000US-207454P.  
PR 09-JUN-2000: 2000US-211314P.  
XX 18-JUL-2000: 2000US-219007P.  
PR 13-DEC-2000: 2000US-255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
XX WPI: 2001-662795/76.  
DR  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1: Page 6318: 11750pp: English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
XX Sequence 2277 BP: 644 A: 568 C: 653 G: 412 T: 0 other:  
SO

Query Match 98.0%; Score 1621.4; DB 23; Length 2277;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1628; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 GGGGGCTCTCGAGACACGGCGGTGTCGTGTGGCCATGGCCCTACCTGATTA 66  
DB 91 GCGCGTCTCTCGAGACACGGCGGTGTCGTGTGGCCATGGCCCTACCTGATTA 150  
QY 67 GTGGGGGACACCTCTACCTGTCGAGACGAGGACTACAGACAGCTTCAACTGCG 126  
DB 151 GTGGGGGACACCTCTACCTGTCGAGACGAGGACTACAGACAGCTTCAACTGCG 210  
QY 127 GAGAGGCTCTACCTACATGACTTTCATCTCCCTGGTACATCCGCTTCACTGAG 186  
DB 211 GAGAGGCTCTACCTACATGACTTTCATCTCCCTGGTACATCCGCTTCACTGAG 270  
QY 187 ACCAGGTGACCTGACTTCTGCTTGACCAAGAAATCACTTAAAGTCCCACTGGTTT 246  
DB 271 ACCAGGTGACCTGACTTCTGCTTGACCAAGAAATCACTTAAAGTCCCACTGGTTT 330

QY 247 CCTCTCCCATGACACAGTCACAGAGGCTGGCATGGCCATAGCAATGGCCTTACAGCGC 306  
DB 331 CCTCTCCCATGACACAGTCACAGAGGCTGGCATGGCCATAGCAATGGCCTTACAGCGC 390  
QY 307 GTAATGGCTTCACTCCACCAACCTGATACCTGATTTCCAGGCCAATGAAAGTTCGGAAG 366  
DB 391 GTAATGGCTTCACTCCACCAACCTGATACCTGATTTCCAGGCCAATGAAAGTTCGGAAG 450  
QY 367 TGAAGAAATATGACAGGACTTTCATACAGACCTGTGTCTTCAGCCCCAAGATTCGGC 426  
DB 451 TGAAGAAATATGACAGGACTTTCATACAGACCTGTGTCTTCAGCCCCAAGATTCGGC 510  
QY 427 TGGCGGATGTTTTTGAAGCCCAAGGCCGATGCTTTCGGGTATGCCAATCACAGACA 486  
DB 511 TGGCGGATGTTTTTGAAGCCCAAGGCCGATGCTTTCGGGTATGCCAATCACAGACA 570  
QY 487 CAGGCGGATGGGAGCCGCTTGTGGGATCATCTCTCCAGGACATTTGATTTCTCA 546  
DB 571 CAGGCGGATGGGAGCCGCTTGTGGGATCATCTCTCCAGGACATTTGATTTCTCA 630  
QY 547 AAGAGGAGAACATGACTGTGTTCTTGAAAGAGTAAATGACAAAGAGGGAAGCTGTGG 606  
DB 631 AAGAGGAGAACATGACTGTGTTCTTGAAAGAGTAAATGACAAAGAGGGAAGCTGTGG 690  
QY 607 TAGCCCCCGACAGCATCACACTGGAAGAGCAATGAAATTTCTGACGCCAGCAAGAACG 666  
DB 691 TAGCCCCCGACAGCATCACACTGGAAGAGCAATGAAATTTCTGACGCCAGCAAGAACG 750  
QY 667 GAAAGTTGCCCATTTGAAATGAAAGATGATGACTTGTGGCCATATTTGCCCGGACAGACC 726  
DB 751 GAAAGTTGCCCATTTGAAATGAAAGATGATGACTTGTGGCCATATTTGCCCGGACAGACC 810  
QY 727 TGAAGAAATATGGGACTACCTACCTCCAAAGATCCCAAGAAACAGCTGTGTGTG 786  
DB 811 TGAAGAAATATGGGACTACCTACCTCCAAAGATCCCAAGAAACAGCTGTGTGTG 870  
QY 787 GGGCACCCTATGGGACTCATGAGATGACAAAGTATAGCTGTGGCTTGTGCCCGACGCTG 846  
DB 871 GGGCACCCTATGGGACTCATGAGATGACAAAGTATAGCTGTGGCTTGTGCCCGACGCTG 930  
QY 847 GTGTGATGTAGTGTGTTTGGACTCTTCCAGGAAATTCATCTTCAGATCAATATGA 906  
DB 931 GTGTGATGTAGTGTGTTTGGACTCTTCCAGGAAATTCATCTTCAGATCAATATGA 990  
QY 907 TCAAGTACATCAAGCAAAATACCTAATCTCAAGTCAATTTGAGGCAATTTGTGCTACTG 966  
DB 991 TCAAGTACATCAAGCAAAATACCTAATCTCAAGTCAATTTGAGGCAATTTGTGCTACTG 1050  
QY 967 CTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGCCCTGCGGGGAGCATGGCAA 1026  
DB 1051 CTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGCCCTGCGGGGAGCATGGCAA 1110  
QY 1027 GTGGCTCCATCTGCAATATCCAGGAAGTGTGGCTGTGGGCGGCCCAAGCAACAGCAG 1086  
DB 1111 GTGGCTCCATCTGCAATATCCAGGAAGTGTGGCTGTGGGCGGCCCAAGCAACAGCAG 1170  
QY 1087 TGTACAAAGTGTATAGTATGACAGGCGCTTGTGGTTCGGGTCAATTTGTGTGGAGAA 1146  
DB 1171 TGTACAAAGTGTATAGTATGACAGGCGCTTGTGGTTCGGGTCAATTTGTGTGGAGAA 1230  
QY 1147 TCCAAATATGTGGTCAATATTTGCGAAAGCCTTGGCCCTTGGGCGCTCCACAGTATGATG 1206  
DB 1231 TCCAAATATGTGGTCAATATTTGCGAAAGCCTTGGCCCTTGGGCGCTCCACAGTATGATG 1290  
QY 1207 GCTCTCTCTGCTGCGACCATGAGGCCCTTGGGAATATCTTTTCCGATGGGATTC 1266  
DB 1291 GCTCTCTCTGCTGCGACCATGAGGCCCTTGGGAATATCTTTTCCGATGGGATTC 1350  
QY 1267 GGCTAAAGAAATATGCGGATATGCTTCTGTGATGCCATGACAAAGCAACCTCACAGACC 1326  
DB 1351 GGCTAAAGAAATATGCGGATATGCTTCTGTGATGCCATGACAAAGCAACCTCACAGACC 1410  
QY 1327 AGAACGATATTTTCACTGACAGTGCACAAATCAAAAGTGCGCCAGGAGAGTGTCTGCTCG 1386





QY	495	ATGGGAGCGCTTGGTGGGACATCATCTCTCCAGGACATTGATTTTCTCAAGAGAG	554
Db	440	ATGGGAGCGCGTGGTGGGACATCATCTCTCCAGGACATTGATTTTCTCAAGAGAG	499
QY	555	GACATACACTGTTCTTGGGAAGATTAATGACAAAGAGGAAAGACTTGGTGGAGCCCC	614
Db	500	GACATACACTGTTCTTGGGAAGATTAATGACAAAGAGGAAAGACTTGGTGGAGCCCC	559
QY	615	CGGACATCACACAGGAAGAGGCGCAATGAAATCTCCACGCGACGACGAAGAGGAAGTTG	674
Db	560	GCGAGCATCACACTGGAAGAGGCGCAATGAAATCTTCCAGCGACGACGAAGAGGAAGTTG	619
QY	675	CCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGGACAGACCTGGAAG	734
Db	620	CCCATTTGTAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGGACAGACCTGGAAG	679
QY	735	AATGGGACATACCCACTAGCTTCCAAAGATGCCAAGAACAAGCTCTGTGTGGGAGCC	794
Db	680	AATGGGACATACCCACTAGCTTCCAAAGATGCCAAGAACAAGCTCTGTGTGGGAGCC	739
QY	795	ATTGGCACTATGAGGATGATGACAGTPTAGGCTGGAGCTTCTCGCCACAGGCTGGTGTGAT	854
Db	740	ATTGGCACTATGAGGATGATGACAGTPTAGGCTGGAGCTTCTCGCCACAGGCTGGTGTGAT	799
QY	855	GTAGTGGTTTTGGACTTCTCCAGGGAATTCATCTTCCAGATCAATATGATCAAGTAC	914
Db	800	GTAGTGGTTTTGGACTTCTCCAGGGAATTCATCTTCCAGATCAATATGATCAAGTAC	859
QY	915	ATCAAAACAAATACCTTAATCTCCAAATCATTTGGAGGCAATGTGTCATCTGTCCAG	974
Db	860	ATCAAAACAAATACCTTAATCTCCAAATCATTTGGAGGCAATGTGTCATCTGTCCAG	919
QY	975	GCCAAAGACCTCATTTGATGACAGGTGTGATATGCCCTCGGGGTGGGATGGAAATGGCTCC	1034
Db	920	GCCAAAGACCTCATTTGATGACAGGTGTGATATGCCCTCGGGGTGGGATGGAAATGGCTCC	979
QY	1035	ATTCGATTTATCCAGGAAGTGTGGCTGTGGCGGCGCCCAAGCAACAGCATGTACAAG	1094
Db	980	ATTCGATTTATCCAGGAAGTGTGGCTGTGGCGGCGCCCAAGCAACAGCATGTACAAG	1039
QY	1095	GTGATAGATGATGACGCGCCCTTGGTGTCCGGTATTGCTGATGAGAGAAATCCAAAT	1154
Db	1040	GTGATAGATGATGACGCGCCCTTGGTGTCCGGTATTGCTGATGAGAGAAATCCAAAT	1099
QY	1155	GTGGGTATPTTGGGAAAGCCTTGGCCCTTGGGGCTCCACAGTCATGATGGGCTCTCTC	1214
Db	1100	GTGGGTATPTTGGGAAAGCCTTGGCCCTTGGGGCTCCACAGTCATGATGGGCTCTCTC	1159
QY	1215	CTGGCTCCACACACTGAGGCCCTGTGTGATAATCTTTTTCGATGGGATCCGCTAAAG	1274
Db	1160	CTGGCTCCACACACTGAGGCCCTGTGTGATAATCTTTTTCGATGGGATCCGCTAAAG	1219
QY	1275	AAATATGGCGTATGGGTTCTCTCGATGCCATGACAGACACCTCAGCAGCCAGACAGA	1334
Db	1220	AAATATGGCGTATGGGTTCTCTCGATGCCATGACAGACACCTCAGCAGCCAGACAGA	1279
QY	1335	TATTTACGTAGCGTGCACAAATCAAAAGTGGGCCAGAGGAGTCTGTGGTGTGTGAGGAC	1394
Db	1280	TATTTACGTAGCGTGCACAAATCAAAAGTGGGCCAGAGGAGTCTGTGGTGTGTGAGGAC	1339
QY	1395	AAAGGCTCAATCCACAATTTGTCCCTTACTGTATGGCTGGAGTCCCAACACTCATGCGAG	1454
Db	1340	AAAGGCTCAATCCACAATTTGTCCCTTACTGTATGGCTGGAGTCCCAACACTCATGCGAG	1399
QY	1455	GACATTTGGTCCCAAGAGCTTGACCCAAAGTCCGAGCCATGATGTACTGTGGGAGCTTAAG	1514
Db	1400	GACATTTGGTCCCAAGAGCTTGACCCAAAGTCCGAGCCATGATGTACTGTGGGAGCTTAAG	1459
QY	1515	TTTTCAGAGCAAGACGTCCTCAGGCCAGGTGGAGAAAGTGGGGTCCA-TAGCTTCATTTGTA	1573
Db	1460	TTTTCAGAGCAAGACGTCCTCAGGCCAGGTGGAGAAAGTGGGGTCCATTTAGCTTCATTTGTA	1519

OY	1574	T-GAAGACGGCGTTTTCGAAAAAGGATCCACAC-ACCTCCTCGGTTTTTTTTCATAA	1631
Dd	1520	TGGAAAGCGGGGCTTCTGAAAAGGATCCACACAACCTCCTCGTTTTTTTTCACATA	1579
OY	1632	AAAGTTT 1638     	
Dd	1580	ACACGTT 1586	
RESULT 5			
XX	ID	AAZ41300	
XX		AAZ41300 standard; CDNA: 1905 BP.	
AC		AAZ41300;	
XX		18-JAN-2000 (first entry)	
DT			
XX		Human normal ovarian tissue derived CDNA 79.	
DE			
XX		Human; ovary; screening; ovarian cancer; treatment; ss.	
KW			
OS		Homo sapiens.	
XX			
PN		DEL9816395-A1.	
PD		07-OCT-1999.	
XX			
PF		03-APR-1998; 98DE-1016395.	
XX			
PR		03-APR-1998; 98DE-1016395.	
XX			
PA		(META-) METAGEN GES GENOMFORSCHUNG MBH.	
PI		Rosenthal A, Specht T, Hinmann B, Schmitt A, Pilarsky C, Dahl E;	
DR		WPL: 1999-552352/47.	
XX			
PT		Nucleic acid sequences potentially useful in diagnosis or therapy of	
PT		ovarian cancer -	
PS		Claim 3; Page 190; 274ppp; German.	
XX			
CC		This invention describes novel nucleic acid sequences that are highly	
CC		expressed in normal ovary tissue. Artificial chromosomes and cosmid	
CC		clones containing the sequences can be used as gene transfer vehicles.	
CC		The sequences can be used to produce DNA fragments containing	
CC		full-length genes. Host cells transformed with the sequences can be used	
CC		to produce polypeptides or polypeptide fragments, which can be used to	
CC		screen phage displays for polypeptides that bind to them, or as tools for	
CC		identifying agents active against ovarian cancer, or to prepare	
CC		medicaments for treating ovarian cancer. The cDNA sequences can be used	
CC		to obtain genomic genes, their promoters, enhancers, silencers, exon	
CC		structures, intron structures and their splice variants. AAZ41222-241324	
CC		represent cDNA sequences derived from normal human ovarian tissue and	
CC		which encode the protein fragments represented in AAY9724-Y59837.	
XX			
SQ		Sequence 1905 BP; 466 A; 489 C; 534 G; 416 T; 0 other:	
Query Match	87.6%;	Score 1449.4;	DB 20; Length 1905;
Best Local Similarity	94.7%;	Pred. No. 0;	
Matches 1552; Conservative	0;	Mismatches 11;	Indels 76; Gaps 22
OY	7	GGGCGGTCCTCGGAACACGCGCGGTGCCTGTGTGGCCATTGGCGCACTACCTGATTAA	66
Dd	319	GCGCGGTCTCTCGGAACACGCGCGCGGTGTCTCTGTGTGGCCATTGGCGCACTACCTGATTAA	378
OY	67	GTGGGGGACAGCTCTACAGTGCCAGACGACGACTCACACAGCAGCATCTTCAACTGCCG	126
Dd	379	GTGGGGGACAGCTCTACAGTGCCAGACGACGACTCACACAGCAGCATCTTCAACTGCCG	438
OY	127	GAGACGGCTCACCTACAAATGACTTTTCTCCCTGGTGATCATGCATTCACCTGACG	186
Dd	439	GAGACGGCTCACCTACAAATGACTTTTCTCCCTGGTGATCATGCATTCACCTGACG	498



```
QY 187 ACCAGGTGACCTGACTTCGTGCTGTGACCAAGAATACTCTTAAGCCCCACTGTTT 246
    |||||||
Db 499 ACCAGGTGACCTGACTTCGTGCTGTGACCAAGAATACTCTTAAGCCCCACTGTTT 558
QY 247 CCTCTCCATGACACAGTTCACAGAGCTGGATGGCCATAGCAATGCGCTTACAGGC 306
    |||||||
Db 559 CCTCTCCATGACACAGTTCACAGAGCTGGATGGCCATAGCAATGCGCTTACAGGC 618
QY 307 GATTTGGCTTCATCCACCAACTGTACACTGTAATTCAGGCCATGAAGTTGGAAAG 366
    |||||||
Db 619 ----- 618
QY 367 TGAAGAAATATGACAGGAGATTCATGACAGACCTGTGCTCTACGCCCAAGAGATCGC 426
    |||||||
Db 619 -----ATGATTCATCAGACCTGTGCTCTACGCCCAAGAGATCGC 663
QY 427 TGGGGATGTTTGTGAGGCCAAGGCCGCGATGTTTGTGGCGGTATCCCAATCAGACACA 486
    |||||||
Db 664 TGGGGATGTTTGTGAGGCCAAGGCCGCGATGTTTGTGGCGGTATCCCAATCAGACACA 723
QY 487 CAGGCCGATGGGAGCCCGCTTGGTGGGATCATCTCTCCAGGACATTTGATTTTCTCA 546
    |||||||
Db 724 CAGGCCGATGGGAGCCCGCTTGGTGGGATCATCTCTCCAGGACATTTGATTTTCTCA 783
QY 547 AAGAGGAGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGCAAGACTTGGTGG 606
    |||||||
Db 784 AAGAGGAGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGCAAGACTTGGTGG 843
QY 607 TAGCCCCCGACAGCATCACACTGTAAGAGGCAAAATTTCTGACGCCAGCAGCAAGAG 666
    |||||||
Db 844 TAGCCCCCGACAGCATCACACTGTAAGAGGCAAAATTTCTGACGCCAGCAGCAAGAG 902
QY 667 GAAAGTTGCCATTTGAATGAAGATGATGAGTGTGGCCATCATTCGCCGACAGACC 726
    |||||||
Db 903 GAAAGTTGCCATTTGAATGAAGATGATGAGTGTGGCCATCATTCGCCGACAGACC 962
QY 727 TGAAGAAGATCGGAGCTACCCACTAGCTCCAAAGATGCCAAAGAACGCTCTGTGTG 786
    |||||||
Db 963 TGAAGAAGATCGGAGCTACCCACTAGCTCCAAAGATGCCAAAGAACGCTCTGTGTG 1022
QY 787 GGGACCCCATTTGGCACTCATGAGATGACAAAGATAGCTGGACCTTGCGCCAGGCTG 846
    |||||||
Db 1023 GGGACCCCATTTGGCACTCATGAGATGACAAAGATAGCTGGACCTTGCGCCAGGCTG 1082
QY 847 GTGTGATGTAGTGTGTTTGGACTCTTCCAGGGAATTCATCTTCCGATCAATATGA 906
    |||||||
Db 1083 GTGTGATGTAGTGTGTTTGGACTCTTCCAGGGAATTCATCTTCCGATCAATATGA 1142
QY 907 TCAAGTACATCAAAAGCAAAATACCTTAATCTCAAGTCAATTGGAGGCAATGTGCTACTG 966
    |||||||
Db 1143 TCAAGTACATCAAAAGCAAAATACCTTAATCTCAAGTCAATTGGAGGCAATGTGCTACTG 1202
QY 967 CTGCCAGGCGCAAGAACTCATTTGATGCAAGGTGTGATGCCCTGGCGGGGGGCGG 1026
    |||||||
Db 1203 CTGCCAGGCGCAAGAACTCATTTGATGCAAGGTGTGATGCCCTGGCGGGGGGCGG 1262
QY 1027 GTGTGCTCATCTGTCATTAATCCAGGAAGTCTGTGCGCTGTGCGCGGCCCCAGCAACAGCAG 1086
    |||||||
Db 1263 GTGTGCTCATCTGTCATTAATCCAGGAAGTCTGTGCGCTGTGCGCGGCCCCAGCAACAGCAG 1322
QY 1087 TGTACAAAGTGTATGATGATGACAGCGGCTTGGTGTTCGGGTCATTTGATGAGGAGAA 1146
    |||||||
Db 1323 TGTACAAAGTGTATGATGATGACAGCGGCTTGGTGTTCGGGTCATTTGATGAGGAGAA 1382
QY 1147 TCCAAAATGTGGTCAATATTGGGAAAGCCTTGGCCCTTGGGCGCTCCCAAGTCATGATGG 1206
    |||||||
Db 1383 TCCAAAATGTGGTCAATATTGGGAAAGCCTTGGCCCTTGGGCGCTCCCAAGTCATGATGG 1442
QY 1207 GCTCTCTCTGGTGCACACACTGAGAGCCCTGTGTAATACCTTTTCCGATGGGATCC 1266
    |||||||
Db 1443 GCTCTCTCTGGTGCACACACTGAGAGCCCTGTGTAATACCTTTTCCGATGGGATCC 1502
```

```
QY 1267 GCCTAAAGAAATATTCGCGGATGTTCTCTGATGCATGACAGCAAGCAACCTCAGCAGCC 1326
    |||||||
Db 1503 GCCTAAAGAAATATTCGCGGATGTTCTCTGATGCATGACAGCAAGCAACCTCAGCAGCC 1562
QY 1327 AGAAGAGATATTTTCAGTGAAGCTGACAAAAATCAAGTGGCCAGGAGAGTCTGTGCTTG 1386
    |||||||
Db 1563 AGAAGAGATATTTTCAGTGAAGCTGACAAAAATCAAGTGGCCAGGAGAGTCTGTGCTTG 1622
QY 1387 TGCAGAGCAAAAGGCTCAATCCACAATTTGTCCCTTACCTGATTTGGTGGCATCCACAC 1446
    |||||||
Db 1623 TGCAGAGCAAAAGGCTCAATCCACAATTTGTCCCTTACCTGATTTGGTGGCATCCACAC 1682
QY 1447 CATGCCAGGACATTTGTCGCCAAGAGCTTGACCCAGTCCGAGGCATGATGTACTGTGGG 1506
    |||||||
Db 1683 CATGCCAGGACATTTGTCGCCAAGAGCTTGACCCAGTCCGAGGCATGATGTACTGTGGG 1742
QY 1507 AGCTTAAGTTTGAAGAGAGAGAGCTCCGAGCCAGGTGGAAAGTGGCGTCCATAGGCTCC 1566
    |||||||
Db 1743 AGCTTAAGTTTGAAGAGAGAGAGCTCCGAGCCAGGTGGAAAGTGGCGTCCATAGGCTCC 1802
QY 1567 ATTCTATGAGAGAGCGCTTTCTGAAAAGGATCCAGCAGCACCTCGGTGTTTTTTTTT 1626
    |||||||
Db 1803 ATTCTATGAGAGAGCGCTTTCTGAAAAGGATCCAGCAGCACCTCGGTGTTTTTTTTT 1862
QY 1627 CAATAAAGCTTTGAAAAGA 1645
    |||||||
Db 1863 CAATAAAGCTTTGAAAAGA 1881

RESULT 6
AA003541
ID AA003541 strand:rd; DNA; 1620 Bp.
XX
AC AA003541:
XX
DT 31-AUG-1990 (first entry)
XX
DE Chinese hamster IMPDH.
XX
KW Inosine 5'-mono-phosphate dehydrogenase; hepatomas;
KM guanosine monophosphate; ss.
XX
OS Cricetus sp.
XX
PN WO9001545-A.
XX
PD 22-FEB-1990.
XX
PF 02-AUG-1989; 89WO-0000344.
XX
PR 12-AUG-1988; 88US-0232302.
XX
PA (ARCH-) ARCH. DEV. CORP.
XX
PI Collart FR, Huberman E;
XX
DR WPI: 1990-083504/11.
XX
DR N-PSDB: AA003541.
XX
PT DNA encoding eukaryotic inosine 5'-mono-phosphate dehydrogenase -
PT esp. hepatomas; and to produce guanosine monophosphate.
XX
PS Claim 8; Fig 2; S1pp: English.
XX
CC The sequence can be used to quantitatively detect IMPDH encoding
CC DNA or RNA by hybridisation, eg in normal and malignant cells,
CC esp. hepatomas.
CC See also AA003540.
XX
SQ Sequence 1620 Bp; 423 A; 369 C; 421 G; 406 T; 1 other:
XX

Query Match 79.8%; Score 1320.2; DB 11; Length 1620;
```

Best Local Similarity 88.8%; Pred. No. 0;	
Matches 1441; Conservative 0; Mismatches 178; Indels 4; Gaps 1	
OY	23 CACGCGCGGTCTCTGTGTGGCCATGGCCGACTACCTGATTAGTGGGGCAGCTCTTA 82
Db	1 CACGGCTGCTCTCTCTGTTGGCCATGGCGACTACCTGATTAGGGGAGCACAATCTTA 60
OY	83 CGTGCAGCGAGGCACTCACAGCAGAGCGTCTTCACTCGGGAGCGGGCTCACTTA 142
Db	61 CTTGCCGACGAGGGGCTACAGCGCAGAGCTCTTCAACTCGCGGGATGGCTTCACTTA 120
OY	143 CAATGACTTTCTCATTTCTCCCTGGGTACATGCACTTCACTGCAGACCAAGTGGACTGAC 202
Db	121 CAACGATTTTCTCATTTCTTCTGGGTATATGCACTTCACTGCGCGACCAAGTGGATTGGAC 180
OY	203 TTTCTGCTGACCAAGAAAATCAGCTTTAAGACCCCACTGGTTTCTCTCCCATGGAGAC 262
Db	181 CTTCTGCTCAACTAAGAAATACACCTTGAGACCCCACTGGTTTCTCTCACTTAGGACAC 240
OY	263 AGTCACAGAGCGTGGGATGGCCATAGCAATGGCGCTTTACAGCGGTATTTGGCTTCATCA 322
Db	241 TGTCAAGAGGGGTGGAAATGGCCATTGCATTTGGGCTTTACAGAGGTATTTGGCTTCATCA 300
OY	323 CCACAACTGTACACTGGAATTCGAGGCCAATGAAGTTCCGAAAAGTAAAGAAATATGAACA 382
Db	301 CCACAACTGTACACTGGAATTCGAGGCCAATGAAGTTCCGAAAAGTAAAGAAATATGAACA 360
OY	383 GGGATTTCATCAGACACCTCGTGCTCTCAGGCCCAAGGATTCGGCTCGGGATGTTTTTGA 442
Db	361 GGGATTTCATCAGATCCTGTAGVCTTATGGCCCAAGGATTCGTTGAGGGATGTTTTTGA 420
OY	443 GGGCAAGCGCCGCGCATGTTTCTGCGGTATCCCAATCACAGACACAGCGCGGATGGGAG 502
Db	421 AGCCAAAGCGCAAGCATGGCTTCTGTGTATCCCATCACAGATACAGCGCGGATGGGAG 480
OY	503 CGGCTTGGTGGCATATATCTCCCGAGGACATGATTTTCTCAAGAGGAGGAACATGA 562
Db	481 TCGACTGGTGGCATATATTTCTTCAAGGGATTTGATTTTCTCAAGAGGAGGACATGA 540
OY	563 CTGTTTCTTGGAAAGAGATATATGACAAAGAGGGAAGACTTGGTGTATGACCCCGCGACAT 622
Db	541 CGGTTTCTTGGAGAGATCATGACAAAGAGGGAAGATTTGGTGTGGCCCTGACAGCAT 600
OY	623 CACACTGAAGGAGGCAATATGAATTTCTGACGGCAGCAGAAAGGAAATTTGCCATTCT 682
Db	601 CACTCTGAAGGAGGCAATATGAATTTCTGACGGCAGCAGTAAAGGAAAGTTGCCATTCT 660
OY	683 AAATGAAGATGATGACCTTGCGCCATCATTTGGCCGAGACAGCTCAAGAAAGATGGGA 742
Db	661 GAATGAAGATGATGACCTGTACCATCATTTCTCGAGACAGACTGAAGAAAGATCGTGA 720
OY	743 CTACCCACTAGCCTCCAAAGATGCCAAGAAACAGCTGTCTGTGGGGCAGCCATTGGCAC 802
Db	721 TTACCCATTGGCTTCCAAAGATGCCAAGAAAGGACTATATGTGGGGCAGCCATTGGTAC 780
OY	803 TCATGAGATGACAACTATATAGCTGACCTTGCGCCAGCGCTGAGTGTAGTGGT 862
Db	781 TCATGAGATGACAACTATATAGCTGACCTTGCGCTTGTCTGGTGTGTAGTGGT 840
OY	863 TTTTGACTTTCCACAGGAATTCATCTTCCAGATCAATATATGATCAAGTATCAAGA 922
Db	841 TTTTGACTTTCCACAGGAATTCATCTTCCAAATCAATATATGATCAATATCAATGAAGA 900
OY	923 CAATATCCCTAATCTCCAAATCATTTGGAGGCAATGTGTACTGTCTGCCAGGCCAAGAA 982
Db	901 GAAATACCCCAATCTCCAAATCATTTGGAGGCAATGTATGTACTGTCTGTCAACCAAGAA 960
OY	983 CCGCATTTGATGACAGGTGTGATGCCCGCGGGGCGGATGGGAAGTGGCTCAATTCAT 1042
Db	961 CCGCATTTGATGACAGGTGTGATGCTCTTGGAGTGTGGCATGGGCTGTCTCATTCAT 1020
OY	1043 TATTCACGAGAGTCTGAGCTGTGGCGGCCCAAGCAACAGCATGTACAGGTTGATGA 1102
Db	1021 TATTCACGAGAGTCTGAGCTGTGGCGGCCCAAGCAACAGCATGTACAGGTTGATGA 1000

Db	1021	TATCAGAGAAAGTGTGTGGCCGTGTGTCGGCCGCCCAACAGCAACAGAGTGTCAAGGTTTCTGA	1080
Qy	1103	GTATGACAGGCGCTTTGGTGTTCGGCTATCTCTCATGTGAGAGATCCAAAATGTGGTCA	1162
Db	1081	GTATGTCGGCGCTTGTGTGTCTTGTATTGTCTGATGTAGAGAAATCCAAAATGTGGTCA	1140
Qy	1163	TATTGCGAAGACCTTGGCCCTTGGGGCCCTCCACACTGCATGATGGGCTCTCCGCGGTGC	1222
Db	1141	TATTGCCAAGCTTTGGCTCTTGGACCTTCTACAGTCATGATGGGCTCTCCCTTGGCTGC	1200
Qy	1223	CACACTAGAGCCCTGTGTGATFACCTCTTTTCGATGGAGATCCGGCTAAAGAAATATTCG	1282
Db	1201	CACCACCGAAGCCCTGTGTGTAGTACTTCTTCTCAGATGGGATCCGGCTAAAGAAATATTCG	1260
Qy	1283	CGGTATGGGTTCTCTGCATCCATGTGACAAGACCTCAGCAGCAGAAACAGATTTTTCAG	1342
Db	1261	TGTTATGGGTTCTTGTATGCCATGTGACAAGCATCTCAGCAGCAGAAACGATTTTTCAG	1320
Qy	1343	TGAAGCTGACAAAATCAAAATGTGGCCCAAGGAGTGTGTGTGTGCGAGCAAAAGGTC	1402
Db	1321	TGAAGCTGACAAAATCAAAATGTGGCCCAAGGAGTGTGTGTGTGCGAGCAAAAGGTC	1380
Qy	1403	AATCCACAAATTTGTCCCTTACCTGATTCGTGCGATCCAAACACTCATGCCAGGACATTTGG	1462
Db	1381	TATCCACAAAGTTCGTCCCTTATTTGATGTGTGCGATCCAGCATTCCTGTCAAGACATTTGG	1440
Qy	1463	TGCCAAGAGCTTGAACCCCAAGTCCAGCCCATGATGTACTGTGGGAGCTTAAAGTTTGAGA	1522
Db	1441	TGCCAAGAGTTTAAACCAAGTCCAGCCCATGATGTACTGTGGGAGCTTAAAGTTTGAGA	1500
Qy	1523	GAGAACGTCTCCAGCCAGGTGGAAGTGGCGCTCATACCTCCATCTGTGTGGAAGCG	1582
Db	1501	GAGAACATCTCCAGCTAGGTGGAAGTGGGTGTCCACACCTCTCATTTGTGTGGAAGCG	1560
Qy	1583	GCTTTTCTGAAGAGGATCCAGCACACCTCTCTGGCTTTTTCATATAAAGTTTGAAG	1642
Db	1561	GCTTTTCTGAAGAGGATCCAGTA----TATGCTTGAATTTTTCATATAAAGTTTGAAG	1616
Qy	1643	AGA 1645	
Db	1617	AAA 1619	
RESULT 7			
AAF21634			
ID	AAF21634 standard; DNA: 2019 BP.		
XX			
AC	AAF21634:		
XX			
XX	27-MAR-2001 (first entry)		
DE			
XX			
Human breast and ovarian cancer associated antigen gene SEQ ID 21.			
KW	Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;		
KW	neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;		
KW	antidiabetic; antinflammatory; antidiacer; vulnerary; anticonvulsant;		
KW	antibacterial; antifungal; antiparasitic; cardiant; immune disorder;		
KW	Addison's disease; allergy; autoimmune haemolytic anaemia;		
KW	autoimmune thyroiditis; diabetes mellitus; Crohn's disease;		
KW	multiple sclerosis; rheumatoid arthritis; ulcerative colitis;		
KW	cardiovascular disorder; wound healing; neurological disease; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO20005173-A1.		
XX			
PD	21-SEP-2000.		
XX			
PF	08-MAR-2000; 2000MO-US05881.		
XX			
PR	12-MAR-1999; 99US-0124270.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		

XX Rosen CA, Ruben SM;  
 PI  
 XX  
 DR WPI: 2000-611515/58.  
 DR P-PSDB: AAB58731.  
 XX  
 PT New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention,  
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
 PT disorders and neurological diseases -  
 PS  
 PS  
 XX Claim 1: Page 496; 1299pp; English.  
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC of antagonist sequences exhibit cytostatic; immunosuppressive;  
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
 CC antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;  
 CC antibacterial; antifungal; antiparasitic and cardiac activity. The  
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,  
 CC particularly breast and ovarian cancer. The nucleic acid sequences,  
 CC proteins, agonists and antagonists may also be used in the diagnosis,  
 CC prevention and treatment of immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; cardiovascular disorders such as  
 CC myocardial ischaemia; wound healing; neurological diseases such as  
 CC cerebral anoxia and epilepsy; and infectious diseases.  
 CC  
 XX Sequence 2019 BP: 512 A: 483 C: 538 G: 476 T: 10 other:  
 SQ  
 Query Match 74.8%; Score 1236.8; DB 21; Length 2019;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 1263; Conservative 2; Mismatches 9; Gaps 2;

Db 1176 TGGATGTAGTGGTTTGGACCTCTCCAGGAAATTCATCTTCAGATCAATATGATCA 1235  
 Qy 910 AGTACATCAAGCAAAATACCCATCTCCAAAGTCAATGGAGCATTGGTCACTGCTG 969  
 Db 1236 AGTACATCAAGCAAAATACCCATCTCCAAAGTCAATGGAGCATTGGTCACTGCTG 1295  
 Qy 970 CCCAGGCCAAGAACCTCATTTGATGCAGGTGTGATGCCCTGCGGGTGGCATGGAAAGT 1029  
 Db 1296 CCCAGGCCAAGAACCTCATTTGATGCAGGTGTGATGCCCTGCGGGTGGCATGGAAAGT 1355  
 Qy 1030 GCTTCATCTGATTTCCAGCAAGGTGCTGCGGGTGGCATGGAAAGT 1089  
 Db 1356 GCTTCATCTGATTTCCAGCAAGGTGCTGCGGGTGGCATGGAAAGT 1415  
 Qy 1090 ACAAGGTGTATGAGTATGACAGGGCTTTGGTTCCTGCTCATTTGATGAGCAATCC 1149  
 Db 1416 ACAAGGTGTATGAGTATGACAGGGCTTTGGTTCCTGCTCATTTGATGAGCAATCC 1475  
 Qy 1150 AAAATGTGGGTCTATTTGGCAAGCCCTTGCGCCCTTGCGGCTCCAGATCATGATGGGCT 1209  
 Db 1476 AAAATGTGGGTCTATTTGGCAAGCCCTTGCGCCCTTGCGGCTCCAGATCATGATGGGCT 1534  
 Qy 1210 CTCCTCTGCTGCTGACACCTGACAGCCCTTGATGATCTTTTCCGATGGATCCGGC 1269  
 Db 1535 CTCCTCTGCTGCTGACACCTGACAGCCCTTGATGATCTTTTCCGATGGATCCGGC 1594  
 Qy 1270 TAAAGAAATATCCGGGTATGGGCTTCTGATGCCATGGACAAGACCTTCAGCAGCAGA 1329  
 Db 1595 TAAAGAAATATCCGGGTATGGGCTTCTGATGCCATGGACAAGACCTTCAGCAGCAGA 1654  
 Qy 1330 ACAGATATTTCACTGAAGCTGACAAATCAAGTGCCAGGAGGTGTGTGTGTGTG 1389  
 Db 1655 ACAGATATTTCACTGAAGCTGACAAATCAAGTGCCAGGAGGTGTGTGTGTGTG 1714  
 Qy 1390 AGCACAAGGCTCAATCCAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1449  
 Db 1715 AGCACAAGGCTCAATCCAAATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1774  
 Qy 1450 GCCAGACATTTGCTGCAAGAGCTTGAACCCAGTCCGAGCATGATGTACTGTGGGAGC 1509  
 Db 1775 GCCAGACATTTGCTGCAAGAGCTTGAACCCAGTCCGAGCATGATGTACTGTGGGAGC 1834  
 Qy 1510 TTAAAGTTGAGAGAGAGAGCTCTCAGCCAGGTGAGAGGTGCTCCATAGCTCTCAT 1569  
 Db 1835 TTAAAGTTGAGAGAGAGAGCTCTCAGCCAGGTGAGAGGTGCTCCATAGCTCTCAT 1894  
 Qy 1570 CGTATGAGAGAGAGAGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCAA 1629  
 Db 1895 CGTATGAGAGAGAGAGCTTTTCTGAAAAGGATCCAGACACCTCTCGGTTTTTTTCAA 1954  
 Qy 1630 TAAAGTTTGAAGAAGA 1645  
 Db 1955 TAAAGTTTGAAGAAGA 1970

RESULT 8  
 ABR34910  
 ID ABR34910 standard; cDNA: 1157 BP.  
 XX  
 AC ABR34910;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 XX Human cDNA encoding secreted protein #48.  
 DE  
 XX Human: secreted protein; gene: ss; nutritional supplement: haemophilia;  
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumor;  
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;  
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
 KW tissue regeneration; wound healing; burn; hematopoiesis;  
 KW myeloid cell deficiency; lymphoid cell deficiency.

XX Homo sapiens.  
OS  
XX MO200177288-A2.  
PN  
XX 18-OCT-2001.  
PD  
XX 29-MAR-2001; 2001WO-US10224.  
PE  
XX 06-APR-2000; 2000US-195582P.  
PR  
XX (GENEY ) GENETICS INST INC.  
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;  
PI Gulukota K, Graham JR;  
XX  
XX WPI: 2002-179321/23.  
DR  
XX  
XX Five hundred and ninety two polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT -  
XX  
XX Claim 1: Page 93; 372pp; English.  
PS  
XX The invention relates to 592 polynucleotides which have been derived from  
XX a variety of human tissue sources and which encode novel secreted  
XX proteins. The polynucleotides can be used as probes for the  
XX identification and isolation of full length cDNA and genomic DNA. The  
XX polynucleotides and proteins can also be used as nutritional supplements.  
XX The proteins are useful in the treatment of various immune deficiencies  
XX and disorders such as viral infections, bacterial infections, fungal  
XX infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
XX sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
XX and conditions (e.g. asthma). They are also useful for treating  
XX neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
XX disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
XX inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
XX useful for tissue regeneration, for wound healing and in the treatment of  
XX burns, incisions and ulcers. The proteins are also useful for regulating  
XX haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
XX Sequences ABK34863-ABK35454 represent polynucleotides of the invention.  
XX  
XX Sequence 1157 BP; 297 A; 270 C; 326 G; 264 T; 0 other;  
SQ  
Query Match 69.0%; Score 1141.6; DB 24; Length 1157;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1147; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 489 GGGCGGATGGGAGCGCTTGGGCAATCTCCTCCAGGCAATGATTTTTCAAA 548  
DB 1 GGGCGGATGGGAGCGCTTGGGCAATCTCCTCCAGGCAATGATTTTTCAAA 60  
QY 549 GAGGAGCAATGATGCTTCTTGGAAAGATATGCAAGAGGGAAGGCTGGGTGTA 608  
DB 61 GAGGAGCAATGATGCTTCTTGGAAAGATATGCAAGAGGGAAGGCTGGGTGTA 120  
QY 609 GCGCCCGGCGATCATCACTGAAGAGGCAATGAATTTCTGAGCGCAAGAGAAGGGA 668  
DB 121 GCGCCCGGCGATCATCACTGAAGAGGCAATGAATTTCTGAGCGCAAGAGAAGGGA 180  
QY 669 AAGTTGCCCATTTTAATGAAGATGATGAGCTTGTGGCCATTCATTTGCCGAGACCTTG 728  
DB 181 AAGTTGCCCATTTTAATGAAGATGATGAGCTTGTGGCCATTCATTTGCCGAGACCTTG 240  
QY 729 AAGAAGAATGGGACTACCACTAGAGCTCCAAAGATGCCAAGAAAGAGCTGTGTGGG 788  
DB 241 AAGAAGAATGGGACTACCACTAGAGCTCCAAAGATGCCAAGAAAGAGCTGTGTGGG 300  
QY 789 GCAGGCAATGGCACTATGAGATGAGCAAGTATAGGCTGGATGGCTGGCCGAGGCTGGT 848  
DB 301 GCAGGCAATGGCACTATGAGATGAGCAAGTATAGGCTGGATGGCTGGCCGAGGCTGGT 360

QY 849 GTGATGTAGTGTGTTTGGACTCTTCCAGGAAATTCATCTTCACATCAATATGATC 908  
DB 361 GTGATGTAGTGTGTTTGGACTCTTCCAGGAAATTCATCTTCACATCAATATGATC 420  
QY 909 AAGTACATCAAGACAAATATACCTTAATCTCCAAAGTCATTGGAGCAATGTGTCTACTGCT 968  
DB 421 AAGTACATCAAGACAAATATACCTTAATCTCCAAAGTCATTGGAGCAATGTGTCTACTGCT 480  
QY 969 GCCCAGGCCCAAGAACCTCATTTGATGACAGGTGTGATATGCCCTTGGGGTGGCATGTGGAGT 1028  
DB 481 GCCCAGGCCCAAGAACCTCATTTGATGACAGGTGTGATATGCCCTTGGGGTGGCATGTGGAGT 540  
QY 1029 GGCCTCATCTGCATTTATCCAGAGAGTCTGCGCTGTGGGCGGCCCAAGCAACAGCAGTG 1088  
DB 541 GGCCTCATCTGCATTTACGAGAGAGTCTGCGCTGTGGGCGGCCCAAGCAACAGCAGTG 600  
QY 1089 TACAAGGTGTATGATGATCAAGGCGCTTGGTGTTCGGGTATTCGTATGAGAGAAATC 1148  
DB 601 TACAAGGTGTATGATGATCAAGGCGCTTGGTGTTCGGGTATTCGTATGAGAGAAATC 660  
QY 1149 CAAATGTGGTCATATATTCGAAAGGCTTGGCCCTTGGGGCTCCACAGTCATGATGGGC 1208  
DB 661 CAAATGTGGTCATATATTCGAAAGGCTTGGCCCTTGGGGCTCCACAGTCATGATGGGC 720  
QY 1209 TCTCTCTGCTGGTCCACCACTGAGGCCCTGTGTAATCTTCTTTCCGATGGATCCGG 1268  
DB 721 TCTCTCTGCTGGTCCACCACTGAGGCCCTGTGTAATCTTCTTTCCGATGGATCCGG 780  
QY 1269 CTAAAGAAATATCGCGGTATGGGTTCTCTGATGGCATGAGACAGCCTCAGACGCCAG 1328  
DB 781 CTAAAGAAATATCGCGGTATGGGTTCTCTGATGGCATGAGACAGCCTCAGACGCCAG 840  
QY 1329 AACAGATATTTGAGTGAACCTGACAAATCAAAATGAGGCCAGGAGTGTCTGTCTGTG 1388  
DB 841 AACAGATATTTGAGTGAACCTGACAAATCAAAATGAGGCCAGGAGTGTCTGTCTGTG 900  
QY 1389 CAGGACAAAGGGTCAATCCACAATTTGTCCCTTACTTATGTGTGGCATCCAACTCA 1448  
DB 901 CAGGACAAAGGGTCAATCCACAATTTGTCCCTTACTTATGTGTGGCATCCAACTCA 960  
QY 1449 TGGCAGACATTTGGTCCCAAGAGCTTGAACCAAGTCCAGGACATGATTAAGTCTGGAGAG 1508  
DB 961 TGGCAGACATTTGGTCCCAAGAGCTTGAACCAAGTCCAGGACATGATTAAGTCTGGAGAG 1020  
QY 1509 CTTAAAGTTTGAGAAAGAAAGTCTCAGGCCAGGTGGAAGGTGGCTGCATAGCCTCAT 1568  
DB 1021 CTTAAAGTTTGAGAAAGAAAGTCTCAGGCCAGGTGGAAGGTGGCTGCATAGCCTCAT 1080  
QY 1569 TCGTATGAGAAAGCGCTTTTTCGAAAAGGATTCACACACCTCTCGGTTTTTTTTCGA 1628  
DB 1081 TCGTATGAGAAAGCGCTTTTTCGAAAAGGATTCACACACCTCTCGGTTTTTTTTCGA 1140  
QY 1629 ATAAAGTTTAGAAG 1644  
DB 1141 ATAAAGTTTAGAAG 1156  
RESULT 9  
ABK83782  
ID ABK83782 standard; cDNA: 2858 BP.  
XX  
XX ABK83782;  
XX  
XX 14-AUG-2002 (first entry)  
XX  
DE Human cDNA differentially expressed in granulocytic cells #353.  
XX  
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;  
XX viral infection; parasitic infection; protozoal infection;  
XX fungal infection; sterile inflammatory disease; psoriasis;  
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;  
XX adult respiratory distress syndrome; inflammatory bowel disease;



```
||||| 1730 TGGCCCTTGAGCCTCCACAGTATCATGGCTCCCTGTGGCCGCCACTACGAGGCC 1789
1237 CCGGTGAATACCTCTTTTCCGATGGATCCGGCTAAGAAATATCGGGTATGGCTTC 1296
1790 CTGGGCAAGTACTTCTTCTAGACGGGTCGGCTCAAGAAATACGGGGCATGGCTCAC 1849
1297 TCGATGCCATGACAAGCAACCTCAGCAGCCAGAAAGATATTTCACTGAAGCTGACAAA 1356
1850 TGGATCCATGAGAGACAGCAGCCAGAAAGATATCTTCAGCGAGGGGATAAAG 1909
1357 TCAAAATGGCCAGGAGGTCTGTGCTGTGAGAGCAAAAGGCTCAATCCAAATTTG 1416
1910 TCAAGATCCACAGGGGTCTGGGCTCCATCCAGGACAAAGGATCCAGAAATTCG 1969
1417 TCCCTTACCTGATTTGCTGATCCATCCACTCATGCCAGATTTGGTCCAAAGCTTGA 1476
1970 TGGCCATCTCATAGCAGGATCCAAACGGCTGCCAGGATATCGGGGCCCGAGCTGT 2029
1477 CCCAAGTCCGAGCATGATGTACTCTGGGGAGCTTAAGTTGAGAAGAACGCTCTCAG 1536
2030 CTGTCTTCCGATCATGATGATGACTCAGAGAGCTCAAGTTTGAGAAGCGGACATGTGC 2089
1537 CCCAGTGAAGGTGGCGTCCATAGCCTCATTCGTATGAGAAGCGGCTTTTCTGA 1592
2090 CCCAGTGAAGGTGGTGTCTCATGCGCTGCACCTTACGAAAGCGGCTGTACTGA 2145

RESULT 10
ABN95621
ID ABN95621 standard: DNA: 2858 BP.
XX
XX ABN95621:
XX
XX 13-AUG-2002 (first entry)
XX
XX Gene #2119 used to diagnose liver cancer.
XX
XX Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumor; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
XX
XX WO200229103-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001: 2001WO-US30589.
XX
XX 02-OCT-2000: 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Petes-Da-Silva S, Vockley JG:
XX
XX WPI: 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer,
XX hepatocellular carcinoma or metastatic liver tumor in a patient,
XX involves detecting the level of expression of two or more genes in a
XX liver tissue sample.
XX
XX Claim 1: SEQ ID NO 2119: 298bp: English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
```

```
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://pub/published_pcl_sequences.
XX
XX SQ
XX Sequence 2858 BP: 545 A; 923 C; 860 G; 530 T; 0 other:
XX
XX Query Match 58.1%; Score 960.8; DB 24; Length 2858:
XX Best Local Similarity 76.1%; Pred. No. 2e-283;
XX Matches 1184; Conservative 0; Mismatches 372; Indels 0; Gaps 0:
XX
QY 37 CTGTGTTGGCCATGGCCGACCTGATATTAGTGGGGCAGCTCTACGTGCCAGACGACG 96
DB 590 CAGGGCTACCCATGGGGGACTACTGATGACGGCGCCACCGGTATACGGCCGAGATG 649
QY 97 GACTCAGACAGCAGCTCTTCACTGCGGAGAGGGCTCACCTTACATGACTTTCTCA 156
DB 650 GGCTCACCGCGAGCAGCTCTTCGCCAGCGCCGACCTCACATCAACAGACTTCCTGA 709
QY 157 TTTCTCCGGTATCATGACTTCACTGACAGACAGGATGGACCTGACTTGTGCTGACCA 216
DB 710 TTCTCCAGAGATTCTATGACTTCACTGATGAGTGGACTGACCTGACCTGACCTGACC 769
QY 217 AGAAATCACTCTTAAGACCCGACTGGTTCTCTCCATGACACAGTACAGAGGCTG 276
DB 770 GGAAATCAACGGCTGAAGAGCGCATCTCTCCCGCATGAGCACTGTGACAGAGGCTG 829
QY 277 GGATGGCCATGACATGGCGCTTACAGCGGCTATTTGGCTTACATCCACCAACTGTAC 336
DB 830 ACATGGCCATTCGATGCGCTGATGGAGGATTTGGTTCACTTACCAACCACTGACCC 889
QY 337 CTGAATTCAGGCGCAATGAAGTTGGAAAGTGAAGAAATATGACAGAGGATTCATACAG 396
DB 890 CAGAGTTCCAGGCGCAATGAAGTACGACAGGTCAAGAGACTTTGAACAGGGCTTATCACG 949
QY 397 ACCCTGTGCTCTCAGCCGCCCAAGGATCGGTCGGGATGTTTGAAGGCCAAGCCGCGC 456
DB 950 ACCCTGTGCTGCTGAGCCCTCGCACACTGTGGCGGATGTGCGAGGCCAAGATCGCGC 1009
QY 457 ATGCTTCTGGGATATCCCAATTCAGACAGACGCGGATGGGGAGACCCCTTGTGGGCA 516
DB 1010 ATGGCTTCTGTGGCATCCCATCTACGAGAGCGGCACCATGGGCGCAAGCTGTGGGCA 1069
QY 517 TCATCTCTCCAGAGCATGATTTCTCAAGAGGAGAGACATGACTTCTTCTTGAAG 576
DB 1070 TGTCTACCTCCGAGACATGCACTTCTTCTGAGAGAGACCAACACACCTCTCAGTG 1129
QY 577 AGATAATGACAAAGAGGAGAGACTGTGTAGCCGCCCGCAGCATCAGATGAGAGAG 636
DB 1130 AGGTGATGAGCGCCAGATTAAGCTGTGTGGCTCCACAGAGTGTGAGCTTGAAGAG 1189
QY 637 CAATGAATTTCTGACGCCGACAGAGGAAAGTTGCCCATTTGAAATGAAGATGATG 696
DB 1190 CAATGAGATCTCTGACGCGCTAGCAAGAAAGGAGTGTGCTATCTGATGATGATG 1249
QY 697 AGCTTGGGCATCATTTGGCCGAGACCTGAGAGAAAGTGGGACTGCCACTGAGAGATG 756
DB 1250 AGCTGTGGCATCATTCGCCCGACCGACCTGAGAGAAATTCAGACTACCTCTTGCCCT 1309
QY 757 CCAAGATGCCAAGAAACAGCTGCTGTGTGGGCGACCATTTGGCACTCATGAGAGATG 816
DB 1310 CCAAGATTTCCAGAGAGAGCTGCTGTGTGGGCGACCTGTGGGCGACCGTGAAGATG 1369
QY 817 AGTATGCGTGAAGCTTGTGCGCCAGCGCTGTGTGATGATGATGATGATGATGATG 876
DB 1370 AATACGCTGTGACCTGTGACCGACGCGGCGGTCTCAGCTCATGACTTCCACTCTCC 1429
QY 877 AGGGAATTCATCTTCCGATCATATATGATCAAGTACAAAGCAATATCCCAATC 936
DB 1430 AAGGAAATTCGCTATTCAGATGCGCATGTGTGATGATGATGATGATGATGATGATG 1489
```

QY	937	TTCCAAGTCATCAGACGGCATATGTGGTGCACGTCTCCCGACGCAAGAAGCTTCATTGATGCAG	998
Db	1490	TTCCAGGGAATTTGGGGGGAAGCGTGGTGACACACGCCCAAGGCCAAGACGTGATTGATGCTG	1549
QY	997	GTTGGCATGCCCCTTCGCGGTTGGGCGATGGGAAGTGGCTTCATCTTCGATTTCCACGAAAGTGC	1058
Db	1550	GTTGGTACGGGCTCGCCGTGGGGCATGGGCTGGGGCTTCATCTGGCATCTCCAGAAAGTGA	1609
QY	1057	TGGCGCTGTGGCGGGCCGCCCAAGCAACACCACTGTACAAAGTGTATGAGTATTCACGGCGCT	1118
Db	1610	TGGCGCTGTGGCTGGCGGCCCAAGGACACTGTGTATCAAGGTGGCTGTAGTATGGCCGGCGCT	1669
QY	1117	TTTGGTGTTCGCGTCAATGTTGTGATGAGAGAAATCCAAAATGTGGGTCAATTTGGCAAAAGCCT	1178
Db	1670	TTTGGTGTGGCCATCATATGCGGATATGGCGGATCCAGAACCGTGGGACACGTGTCTCAAGGCC	1728
QY	1177	TGGCGCTTTGGGGGCTCTCCACAGTATGATGGGCTCTCTCTGGCTGCTTCACCTAGAGGCC	1238
Db	1730	TGGCGCTTTGGAGCCTCCACAGTATGATGGGCTCTCTCTGGCGGCCCTTACGGAGGCC	1789
QY	1237	CTGGTGATACTCTCTTTTCGGATGGGATCGGCTAAAGAAATATCGGTATGGCTCTC	1298
Db	1790	CTGGCGAGTACTCTCTCTCTCAGACGGGGTGGCGCTCAAGAAAGTACCGGGCATGGCTCAC	1849
QY	1297	TTCGATGCCATGTGACACAGCACCTCAGCACGCCAGAACAGATATTTCACTGTAAGCTGACAAA	1358
Db	1850	TGTGATCCCATGTGAGAAAGAGGACGACGACGCCAGAAAGATACTTCACCGGGGGATAAAG	1909
QY	1357	TCAAGGTGGCCCAAGCGAGTGTCTGTGTCTGTGCTGTGAGAGACAAAGGCTCAATCCACAAATTG	1418
Db	1910	TGAAGATTCGCACAGCGTCTCGCGCTCCATCCAGGACAAAGAGATCCATTCGAAGTTG	1969
QY	1417	TTCCTTACCTGATTTGCTGGCATCCAACTCATCATCCAGAGACATTGGTCCAAAGACTTGA	1478
Db	1970	TGCCCTACCTCATATGCGAGCGCATCCAACACGGCTGCCAGGATATCGGGGCTCCGACGCTGT	2028
QY	1477	CCCAGTCCCAAGCCATCATATCTACCTCTGGGAGGCTTAAGTTTGAGAAATAAGTCTCTCAG	1538
Db	2030	CTGTCCCTTTCGTTCCATGATGTACTCAGAGAGCTCAAGTTTGGAAAGCGACCATGTCTGC	2088
QY	1537	CCCAAGTGAAGAGTGGCGCTCCATAGCCTCCATTGATGAGAAGCGGCTTTTCTGA	1592
Db	2090	CCCAAGTGAAGAGTGGCGTGTCTCATGTGCGCTGCACCTTACGAAAGAGCGCGTACTGA	2145
RESULT 11			
ID	AAS86456	standard; cDNA: 3387 BP.	
XX	AAS86456:		
AC	13-FEB-2002	(first entry)	
DT			
XX			
DE	DNA encoding novel human diagnostic protein #22260.		
KM	Human: chromosome mapping; gene mapping; gene therapy	forensic;	
KW	food supplement; medical imaging; diagnostic; genetic	disorder; ss.	
XX	Homo sapiens.		
OS			
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001: 2001MO-US08631.		
XX			
PR	31-MAR-2000: 2000US-0540217.		
PR	23-AUG-2000: 2000US-0649167.		
XX			
PA	(HYSEQ) HYSEQ INC.		
PI	Dermanac RT, Liu C, Tang YT;		

XX	WPI: 2001-693962/73.
DR	P-PsDB: ABG22269.
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PS	Claim 1; SEQ ID No 22260; 103pp; English.
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human
CC	diagnostic coding sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pcr_sequences.
XX	
SO	Sequence 3387 BP: 677 A; 1069 C; 988 G; 653 T; 0 other:
	Query Match      52.3%; Score 865.6; DB 23; Length 3387;
	Best Local Similarity    74.6%; Pred. No. 3.5e-254;
	Matches 1167; Conservative    0; Mismatches 389; Indels    8; Gaps     6;
OY	37 CTGTCTTGGCCATGGCGCCGACTACGTAGATTAGTGGGGGCAGTCCTACAGGCAGACGACG 96
Db	1077 CAGGCGTAGCCATGGCGGAGCATCGATCAGCGCGGCCGACGGCGTAACGTGCCGAGGATG 1136
OY	97 GACTACACAGCACAGACGCTCTTCACTGCGGAGACGGGCTCACTCAATGAATGACTTTTCTCA 156
Db	1137 GGCTACCGCGGAGAGAGCTCTTGGCCAGGCGGACGACTCACTCAACAAGCACTTCTCTCA 1196
OY	157 TTCTCTCCGNGGACATCATGCACTTCACTGCGAGACAGAGTGAGACTTGATTCCTGTGACCA 216
Db	1197 TTTCCTCCAGGATTCATPAGACTTCACTAGCTGATGAGATGGATGGACCTGAGCCCTGACC 1256
OY	217 AGAAATCACTTTTAAGACCCGACCTGTTTCTCTCTCCATGGACACAGTACAGAGGCTG 276
Db	1257 GGAAGATACAGCTGTGAAGAGCGCACATCATCTCCCTCCCAATGGACACTGTGACAGAGCGTG 1316
OY	277 GGATGGCCATAAACAATGGGCGCTTTAACAGCGGCTATTGGCTTCATCCACCACTGATCAAC 336
Db	1317 ACATGGCCATTTCATATGGCTGTGATGGGAGGATTTGGGTTTCATTCACCACTGACCCC 1376
OY	337 CTGAATTTCCAGGCCCAATGAAGTTGGGAAAAGTAAATAATATGAACAGGATTTCATCAGC 396
Db	1377 CAGAATTTCCAGGCCCAATGAAGTAAAGTCAAAGTCAAAACTTTTAAACAGGCGTTTATCAGG 1436
OY	397 ACCCTGTGGTCTCAGCCCCCAAGATTCGCGTGGCG-GATGTTTTTGGAGGCCAAGGCCCGG 455
Db	1437 ACCCTGTGGTCTCAGCCCCCGCTGGCACACTGTGGGCGATGTCTGTGAGGCGCAAGATCCGG 1496
OY	456 CATGTTTCT-GCGGTATTCCAATACACAGACACAGGCCGATGGGAGCCGCTTGTGTGG 514
Db	1497 CATGGCTTCTCGGGCAATCCCATACGTAGAGACCGGACACATGGGACAGCAACTGTGTGG 1556
OY	515 CATCATCTCTCTCAGAGGACATGATTTTCTCAAAAGAGAGAACTGACTGTTCCTTGGCA 574

```
Db 1557 CATGTCACCTCCGAGACATCGACTTTCCTGCTGAGAAAGACACACCACCCCTCCTCAG 1616
QY 575 AGAGATTAATGACAAAGAGGAGACTGTGGTAGCCCCCGCAGCATCACACTGAAGA 634
Db 1617 TGAGGTGATGAGGCCAAGAGATTGAAGTGGTGGCTCCAGCAGGTGTACCTGTAAGA 1676
QY 635 GGCAAATGAATTTCTGCACGCGCAGCAAGAAAGGAAAGTTGCCATTTGTAATGAAGATGA 694
Db 1677 GGCAATGATGATCTCGACAGCGTAGCAAGAAAGGAAGTGGCTATCGTCAATGATTTGCCA 1736
QY 695 TGAGCTTGCGGCATCATTTGCCCGGACAGACTGAAAGAAATCGGGATACCCTATAC 754
Db 1737 TGAGCTGGGGCATCATCTGCCCCGACCCGACCTGGAAGAGACCGAGCTTACCTTGCG 1796
QY 755 CTCGAAAGATGCGCAAGAAACAGCTGCTGTGGGGCAGCATTTGGCAGCATGAGAGATGA 814
Db 1797 CTCGAAAGATTTCCACGCCGACAGCTCTCTGTGGGGCAGCTGTGGACACCCCGTAGAGAA 1856
QY 815 CAAATATAGGCTGAGACTTGTGCCCCAGGCTGGTGTGATGATGTTTGGACTCTTC 874
Db 1857 CAAATACCTCTGAGACTGCTCACCAGCGGGCGTGCAGCTCATAGTCTTGAGACTGTC 1916
QY 875 CCAGGGAATTTCCATCTTCCAG-ATCATATATGATCAAGTACATCAAGACAAATACCTTA 933
Db 1917 CCAGGGAATTTGCTGTATCAGTACCGCCAGGCTGCTTATTTCAACAGAAAGTACCCGC 1976
QY 934 ATCTCCAACT-CATTGGAGGCAATGCTGCTGCTGCCAGGCCAAGAACCTCATTTGAT 992
Db 1977 ACCTCCAGGTGATGAGGGGGGAAACGTGTGACAGCAGCCAGGCCAAGAACCTGATTTGAT 2036
QY 993 GGAGGTGATGATCCCTGCGGGGTGGGATGGAAAGTGGCTCATCTGATTTATCA-AGA 1051
Db 2037 GCTGGTGTGAGCGGGCTGCGCGGTGGCATGGCTGCGGCTCATCTGATCAGACCCAGGA 2096
QY 1052 AGTGCTGGCTGTGGGGGGCCCCCAACAC-AGCAGTGTACAGAGTGTATGAGTATGC 1108
Db 2097 AGTATGAGGCTGTGGTGGCCCCCAGGCGCTGCTGTACCAAGGTAGGCTGAGTATGC 2156
QY 1109 ACGGCGCTTTGGTGTCCGCTCATTTGATGAGAGAAATCAAAATGTGGGTCATATTGC 1168
Db 2157 CCGGCGCTTTGGTGTGCCATCATATACCGATGCGGCGCATTCAGACCGCTGGAGCACTGCT 2216
QY 1169 GAAAGCTTTGGCCCTTTGGGGGCTCCACAGTATGATGGCTCTCTCTGCTGGCCACCA 1228
Db 2217 CAAAGGCTTGGGCTTTGGAGGCTCCACAGTATGATGGGCTCCCTGCGCGCCACTAC 2276
QY 1229 TGAGGCGCTTGGGATACATTTCTTTCCGATGGGATCGGCTAAAGAAATGCGGGTAT 1288
Db 2277 GAGAGCCCTTGGGAGTACTTCTTCTCAAGCGGGTGGGCTCAAGAGTACCGGGGCA 2336
QY 1289 GGGTCTCTGATGCGCATGAGACAGCAGCTCAAGCAGCAGAAAGATATTTTCAGTGAAGC 1348
Db 2337 GGGCTACAGTGCATGCGCATGAGAGACAGCAGCAGCAGAAAGCATTTCTGAGCAAGG 2396
QY 1349 TGCAAAATCAAAAGTGGCCAGGGAAGTGTCTGTGCTGTGCAGAGCAAAAGGCTCAATCA 1408
Db 2397 GGAATAAAGTGAAGATGCGCAGGAGTGTCTGCGGCTTCATTCAGAGCAAAAGGATCA 2456
QY 1409 CAAATTTGCGCTTACCTGATTTGCTGGCATGCCATCAACAACCTCATGCGCAGAGATTTGCCAA 1468
Db 2457 GAAGTTGTGGCTTACCTCATATAGCAGGCAATCAACAGGCTGCGAGGATTTGGGGCCG 2516
QY 1469 GAGCTTGACCAAGATCCGAGCCATGATGATCTGTGGGAGCTTAAGTTTGAGAAGGAAC 1528
Db 2517 CAGCTGTCTGTCTTCTGCGGTGATGATGATCTGAGAGAGCTCAATTTTGAGAGCGGAC 2576
QY 1529 GTCTCAGGCCAGAGTGGAAAGTGGCTGCATAGACCTTCATTCGTATGAGAAGCGGCTTT 1588
Db 2577 CATGTGCGCCAGAAATTTGAGAGGTGTGTCCATGTGCTGCACTTTACGAAAGCGGCTGTA 2636
QY 1589 CTGA 1592
Db 2637 CTGA 2640
```

```
RESULT 12
AAD28929
ID AAD28929 standard: DNA: 1158 BP.
XX
AC AAD28929:
XX
DT 07-MAY-2002 (first entry)
XX
DE Human type II IMPDH-SPTQ variant DNA.
XX
KW Human: inosine 5'-monophosphate dehydrogenase; IMPDH: drug discovery;
KW proliferative-type disease; cancer; imaging methodology; cytostatic;
KW therapy; enzyme; variant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
WO200185952-A2.
XX
PN 15-NOV-2001.
XX
PD 10-MAY-2001: 2001WO-US15457.
XX
PF 10-MAY-2001: 2000US-203448P.
XX
PR 10-MAY-2000: 2000US-203448P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N:
PI Mounavleff JE, Elnsphaer HM, Kish K:
XX
DR WPI: 2002-164105/21.
XX
PT New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
PT having an oligo peptide domain substituted for a subdomain of a
PT wild-type IMPDH polypeptide, useful in drug discovery or for generating
PT antibodies -
XX
PS Claim 54; Fig 12; 161pp; English.
XX
CC The invention relates to modified inosine 5'-monophosphate dehydrogenase
CC (IMPDH; EC 1.1.1.203) polypeptide, comprising an oligo-peptide domain
CC substituted for a subdomain of a wild-type IMPDH polypeptide. The
CC modified IMPDH polypeptides are useful for drug discovery, for
CC therapeutic, diagnostic and prognostic procedures for detecting or
CC quantifying modified IMPDH polypeptides and their corresponding nucleic
CC acids. IMPDH polypeptides are also useful for generating antibodies, as
CC diagnostic and prognostic markers of diseases, as targets for various
CC therapeutic modalities, and to identify and isolate ligands and other
CC agents that bind to modified IMPDH. These antibodies may be used in
CC diagnostic assays, imaging methodologies, therapeutic methods in the
CC management of cancer or other proliferative-type diseases, and in
CC purifying modified IMPDH polypeptides and for isolating related
CC molecules such as wild type and mutant IMPDH polypeptides. The present
CC sequence is human type II IMPDH-SPTQ variant DNA.
XX
SQ Sequence 1158 BP; 279 A; 301 C; 310 G; 268 T; 0 other:
QY
Query Match 49.4%; Score 816.4; DB 24; Length 1158;
Best Local Similarity 98.7%; Pred. No. 2.2e-239;
Matches 823; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```



Db	445	GGAAATTTCATTCTTCCAGATCAATATGATGCAAGTACATCAAAAGCAAAATACCACTTAATCTC	504
Qy	939	CAAGTCATTGGAGGCAATGTGTCACCTGCTGCCAGGCCCCAAGACCTTAATGATGCAGT	998
Db	505	CAAGTCATTGGAGGCAATGTGTCACCTGCTGCCAGGCCCCAAGACCTTAATGATGCAGT	564
Qy	999	GTGATGACCCCTGGGGGGGCGCATGGGAAGGGCTCCATGCAATTATCAGGAAGTGGT	1058
Db	565	GTGATGACCCCTGGGGGGGCGCATGGGAAGGGCTCCATGCAATTATCAGGAAGTGGT	624
Qy	1059	GCCGTGTGGGGGCGGCCCAAGCAGACAGCAGTGTACAAAGTGTAGATATCCACGGCGCTTT	1118
Db	625	GCCGTGTGGGGGCGGCCCAAGCAGACAGCAGTGTACAAAGTGTAGATATCCACGGCGCTTT	684
Qy	1119	GGTGTTCGGGTCATGCTGTATGGAGGAAATCCAAATGTGGGTCATATTGSCGAAGCCCTTG	1178
Db	685	GGTGTTCGGGTCATGCTGTATGGAGGAAATCCAAATGTGGGTCATATTGSCGAAGCCCTTG	744
Qy	1179	GCCCTTGGGGCTCCACACATCATGATGGGCTCTCTCTGGCTGCACCACTAGAGCCCTT	1238
Db	745	GCCCTTGGGGGCTCCACACAGTCATGATGGGCTCTCTCTGGCTGCACCACTAGAGCCCTT	804
Qy	1239	GGTCAATACCTTCTTTTCCGATGGGATCCGGGCTAAAGAAATATCCGGGATGTGGTCTCTC	1298
Db	805	GGTCAATACCTTCTTTTCCGATGGGATCCGGGCTAAAGAAATATCCGGGATGTGGTCTCTC	864
Qy	1299	GATGCCATGGAACAAGCACCTCAGCACCCAGAAACAGATATTTGAGTGAACTGTACAAAATC	1358
Db	865	GATGCCATGGAACAAGCACCTCAGCACCCAGAAACAGATATTTGAGTGAACTGTACAAAATC	924
Qy	1359	AAAGTGGCCCCAGAGAGTGTGTGGTGTGGAGAGCAAAAGGGTCAATGCACAAATTTTC	1418
Db	925	AAAGTGGCCCCAGAGAGTGTGTGGTGTGGAGAGCAAAAGGGTCAATGCACAAATTTTC	984
Qy	1419	CTTACTGATTTGCTGGCATCCACAACCTATGCGACAGACATTTGGTCCAGAGCTTGAC	1478
Db	985	CTTACTGATTTGCTGGCATCCACAACCTATGCGACAGACATTTGGTCCAGAGCTTGAC	1044
Qy	1479	CAAGTCGAGACCATGATGTACTCTGGGAGAGCTTAAGTTTGAGAAAGAGACGTCCTCACC	1538
Db	1045	CAAGTCGAGACCATGATGTACTCTGGGAGAGCTTAAGTTTGAGAAAGAGACGTCCTCACC	1104
Qy	1539	CAGTGGAGAGGTGGCGTCCATACCTTCATTTGATGAGAGCGGCTTTCTGCA	1592
Db	1105	CAGTGGAGAGGTGGCGTCCATACCTTCATTTGATGAGAGCGGCTTTCTGCA	1158
RESULT 13			
ID	AAD28927	standard; cDNA: 1158 BP.	
XX	XX	AAD28927:	
XX	XX	07-MAY-2002 (first entry)	
XX	XX	Human type II IMPDH-AGRP variant cDNA.	
XX	XX	Human: inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery; proliferative-type disease; cancer; imaging methodology; cytostatic; therapy; enzyme; variant; ss.	
XX	XX	Homo sapiens.	
XX	XX	Synthetic.	
XX	XX	Key	
XX	XX	CDS	
XX	XX	Location/Qualifiers	
XX	XX	1..1158	
XX	XX	/tag= a	
XX	XX	/product= "Human type II IMPDH-AGRP variant"	
XX	XX	W0200185952-A2.	
XX	XX	15-NOV-2001	

PF	10-MAY-2001 : 2001WO-US15457.
XX	
PR	10-MAY-2000; 2000US-203448P.
XX	
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.
XX	
P1	Krystek SR, Sheriff S, Witmer MR, Hollenbaugh DL, Yan N;
PI	Mouravieff JE, Einspahr HM, Kish K;
XX	
DR	WPI: 2002-164105/21.
XX	
DR	P-P5DB: AAE18181.
XX	
PT	New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
PT	having an oligo-peptide domain substituted for a subdomain of a
PT	wild-type IMPDH polypeptide, useful in drug discovery or for generating
PT	antibodies -
XX	
PS	Claim 54; Fig 14; 161pp; English.
XX	
CC	The invention relates to modified inosine 5'-monophosphate dehydrogenase
CC	(IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain
CC	substituted for a subdomain of a wild-type IMPH polypeptide. The
CC	modified IMPDH polypeptides are a useful for drug discovery, for
CC	therapeutic, diagnostic and prognostic procedures for detecting or
CC	quantifying modified IMPDH polypeptides and their corresponding antibodies,
CC	as diagnostic and prognostic markers of diseases, as targets for various
CC	therapeutic modalities, and to identify and isolate ligands and other
CC	agents that bind to modified IMPDH. These antibodies may be used in
CC	diagnostic assays, imaging methodologies, therapeutic methods in the
CC	management of cancer or other proliferative-type diseases, and in
CC	purifying modified IMPDH polypeptides and for isolating related
CC	molecules such as wild type and mutant IMPDH polypeptides. The present
CC	sequence is human type II IMPDH-AcRP variant cDNA.
XX	
SO	Sequence 1158 BP; 277 A; 300 C; 313 G; 268 T; 0 other;
	Query Match 49.2%; Score 813.2; DB 24; Length 1158;
	Best Local Similarity 98.4%; Pred No. 2,1e+238;
	Matches 821; Conservative 0; Mismatches 13; Indels 0; Gaps 0
OY	739 AAAGATGCCAACAACACTGCTGTGTGGGGCAGCCATTGGCACTCATGAGTGCANG 818
DB	325 AAATATGCGTGCGTCGCTGTGTGGGGCAGCCATTGGCACTCATGAGTGCANG 384
OY	819 TATAGCGTAGATTCTCTCCGCCAGCGCTGTGTGATGTAGTGTTTGGACTCTTCCAG 878
DB	385 TATAGCGTAGATTCTCTCCGCCAGCGCTGTGTGATGTAGTGTTTGGACTCTTCCAG 444
OY	879 GGAATTCATTTCCAGATCAATATGATCAGTACATCAAAGACAATAACCTTAATCTC 938
DB	445 GGAATTCATTTCCAGATCAATATGATCAGTACATCAAAGACAATAACCTTAATCTC 504
OY	939 CAAGCATTTGAGGAANTGTGCTACTGCTGGCCAGGCCAAACCTCATTTATGAGAGT 998
DB	505 CAAGCATTTGAGGAANTGTGCTACTGCTGGCCAGGCCAAACCTCATTTATGAGAGT 564
OY	999 GTGAGTGGCTTGGGGTGGGCATGGGAAGTGGCTTCATTCGATTATCCAGAAGTCTTG 1058
DB	565 GTGAGTGGCTTGGGGTGGGCATGGGAAGTGGCTTCATTCGATTATCCAGAAGTCTTG 624
OY	1059 GCCTTGGGGCGGCCCAAGCACAGCAGTGTCAAGAGTTATGAGTATCAGAGCGCTTT 1118
DB	625 GCCTTGGGGCGGCCCAAGCACAGCAGTGTCAAGAGTTATGAGTATCAGAGCGCTTT 684
OY	1119 GGTGTTCCGGTATATGCTGATGAGAGGAATCCAAAATGTGGGTCAATTTGCGAAAGCTTG 1178
DB	685 GGTGTTCCGGTATATGCTGATGAGAGGAATCCAAAATGTGGGTCAATTTGCGAAAGCTTG 744
OY	1179 GCCCTTTGGGGCCTCCACAGTCATGATGGGCTTCTCTCGCTGGCCACACACTAGAGCCCT 1238
DB	745 GCCCTTTGGGGCCTCCACAGTCATGATGGGCTTCTCTCGCTGGCCACACACTAGAGCCCT 804

```
OY 1239 GGTGAATACCTCTTTCCGATGGATCCGGCTAAAGAAATATCGCGATAGGGTCTCTC 1298
|||||
DB 805 GATGATACCTCTTTCCGATGGATCCGGCTAAAGAAATATCGCGATAGGGTCTCTC 864
OY 1299 GATGCATGACACAGACCTCAGACACGACGACGACGATATTTTCAGTAACTGCAAAATC 1358
|||||
DB 865 GATGCATGACACAGACCTCAGACACGACGACGACGATATTTTCAGTAACTGCAAAATC 924
OY 1359 AAAAGTGGCCAGAGAGTGTCTGTGCTGTCAGAGCAAAAGGCTCAATCCAAATTTGTC 1418
|||||
DB 925 AAAGTGGCCAGAGAGTGTCTGTGCTGTCAGAGCAAAAGGCTCAATCCAAATTTGTC 984
OY 1419 CCTTACCTGATTTGCTGGATCCACACACTCATGCGACGACATTTGGTCCAAAGACTTGACC 1478
|||||
DB 985 CCTTACCTGATTTGCTGGATCCACACACTCATGCGACGACATTTGGTCCAAAGACTTGACC 1044
OY 1479 CAAGTCCGAGCCATGATGATCTGTGGAGCTTAAGTTGAGAGAGAACGTCCTCAGCC 1538
|||||
DB 1045 CAAGTCCGAGCCATGATGATCTGTGGAGCTTAAGTTGAGAGAGAACGTCCTCAGCC 1104
OY 1539 CAGGTGGAAGGTGGCTCCATAGCTCCATTCGTATGAGAGAGGCTTTTCTGA 1592
|||||
DB 1105 CAGGTGGAAGGTGGCTCCATAGCTCCATTCGTATGAGAGAGGCTTTTCTGA 1158

RESULT 14
AAD28928
ID AAD28928 standard; cDNA: 1158 BP.
AC AAD28928;
DT 07-MAY-2002 (first entry)
DE Human type II IMPDH-NSPL variant cDNA.
XX
XX
KM Human; inosine 5'-monophosphate dehydrogenase; IMPDH; drug discovery;
KW proliferative-type disease; cancer; imaging methodology; cytosolic;
KW therapy; enzyme; variant; ss.
XX
OS Homo sapiens.
XX
XX Synthetic.
FH Key Location/Qualifiers
FT 1..1158
FT CDS /tag= a
FT /product= "Human type II IMPDH-NSPL variant"
XX
PN WO200185952-A2.
XX
XX 15-NOV-2001.
XX
XX 10-MAY-2001; 2001WO-US15457.
XX
XX 10-MAY-2000; 2000US-203448P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Krystek SR, Sheriff S, Wilmer MR, Hollenbaugh DL, Yan N;
XX PI Mounavleff JE, Einspahr HM, Kish K;
XX DR WPI: 2002-164105/21.
XX DR P-PSDB: AAE18183.
XX
XX New modified inosine 5'-monophosphate dehydrogenase (IMPDH) polypeptide
XX PT having an oligo-peptide domain substituted for a subdomain of a
XX PT wild-type IMPDH polypeptide, useful in drug discovery or for generating
XX PT antibodies
XX
XX Claim 54: Fig 16; 161pp; English.
XX
XX The invention relates to modified inosine 5'-monophosphate dehydrogenase
XX CC (IMPDH; EC 1.1.1.205) polypeptide, comprising an oligo-peptide domain
XX CC substituted for a subdomain of a wild-type IMPDH polypeptide. The
```

```
CC modified IMPDH polypeptides are useful for drug discovery, for
CC therapeutic, diagnostic and prognostic procedures for detecting or
CC quantifying modified IMPDH polypeptides and their corresponding nucleic
CC acids. IMPDH polypeptides are also useful for generating antibodies, as
CC diagnostic and prognostic markers of diseases, as targets for various
CC therapeutic modalities, and to identify and isolate ligands and other
CC agents that bind to modified IMPDH. These antibodies may be used in
CC diagnostic assays, imaging methodologies, therapeutic methods in the
CC management of cancer or other proliferative-type diseases, and in
CC purifying modified IMPDH polypeptides and for isolating related
CC molecules such as wild type and mutant IMPDH polypeptides. The present
CC sequence is human type II IMPDH-NSPL variant cDNA.
XX
XX Sequence 1158 BP; 279 A; 301 C; 309 G; 269 T; 0 other;
SQ
Query Match 49.0%; Score 810; DB 24; Length 1158;
Best Local Similarity 98.2%; Pred. No. 2e-237; Mismatches 15; Indels 0; Gaps 0;
Matches 819; Conservative 0;
OY 759 AAAGATCCCAAGAAACAGCTGCTGTGTGGGCGAGCCATTGGCAGTCATGAGATGACAG 818
|||||
DB 325 AATATTAACCTTCGCCGTTCTGCTGTGTGGGCGAGCCATTGGCAGTCATGAGATGACAG 384
OY 819 TATAGGCTGGACTTGTCTGCCAGGCTGTGTGATGTAGTGGTTTGGACTCTTCCAG 878
|||||
DB 385 TATAGGCTGGACTTGTCTGCCAGGCTGTGTGATGTAGTGGTTTGGACTCTTCCAG 444
OY 879 GGAAATTCATCTCCAGATTCATATGATCAATGATCAATGATCAATGATCAATGATCA 938
|||||
DB 445 GGAAATTCATCTCCAGATTCATATGATCAATGATCAATGATCAATGATCAATGATCA 504
OY 939 CAAGTCATTGAGAGCAATGTGTCACTGTGCCAGGCGCAAGAACCTCATTTGATGAGGT 998
|||||
DB 505 CAAGTCATTGAGAGCAATGTGTCACTGTGCCAGGCGCAAGAACCTCATTTGATGAGGT 564
OY 999 GTGATGATCCCTCGGCTGGGTCATGGAAGTGGCTCATTTGATTCAGAGAGTGTCTG 1058
|||||
DB 565 GTGATGATCCCTCGGCTGGGTCATGGAAGTGGCTCATTTGATTCAGAGAGAGTGTCTG 624
OY 1059 GCCCTGGGCGGCGCCCAACACAGCAGTGTACAAAGTGTATGAGATATGACAGGCGCTT 1118
|||||
DB 625 GCCCTGGGCGGCGCCCAACACAGCAGTGTACAAAGTGTATGAGATATGACAGGCGCTT 684
OY 1119 GGTGTTCCGCTCATGCTATGAGAGGAATCAAAATGTGGTCAATTTGCGAAAGCCTTG 1178
|||||
DB 685 GGTGTTCCGCTCATGCTATGAGAGGAATCAAAATGTGGTCAATTTGCGAAAGCCTTG 744
OY 1179 GCCCTGGGCGGCTCCACAGTCATGATGGCTCTCTCTGCTGCCACCACTGAGGCCCT 1238
|||||
DB 745 GCCCTGGGCGGCTCCACAGTCATGATGGGCTCTCTCTGCTGCCACCACTGAGGCCCT 804
OY 1239 GGTGAATACCTCTTTCCGATGGATCCGGCTAAAGAAATATCGCGATAGGGTCTCTC 1298
|||||
DB 805 GGTGAATACCTCTTTCCGATGGATCCGGCTAAAGAAATATCGCGATAGGGTCTCTC 864
OY 1299 GATGCATGACACAGACCTCAGACACGACGACGACGATATTTTCAGTAACTGCAAAATC 1358
|||||
DB 865 GATGCATGACACAGACCTCAGACACGACGACGACGATATTTTCAGTAACTGCAAAATC 924
OY 1359 AAAAGTGGCCAGAGAGTGTCTGTGCTGTCAGAGCAAAAGGCTCAATCCAAATTTGTC 1418
|||||
DB 925 AAAAGTGGCCAGAGAGTGTCTGTGCTGTCAGAGCAAAAGGCTCAATCCAAATTTGTC 984
OY 1419 CCTTACCTGATTTGCTGGATCCACACACTCATGCGACGACATTTGGTCCAAAGACTTGACC 1478
|||||
DB 985 CCTTACCTGATTTGCTGGATCCACACACTCATGCGACGACATTTGGTCCAAAGACTTGACC 1044
OY 1479 CAAGTCCGAGCCATGATGATCTGTGGAGCTTAAGTTGAGAGAGAACGTCCTCAGCC 1538
|||||
DB 1045 CAAGTCCGAGCCATGATGATCTGTGGAGCTTAAGTTGAGAGAGAACGTCCTCAGCC 1104
OY 1539 CAGGTGGAAGGTGGCTCCATAGCTCCATTCGTATGAGAGAGGCTTTTCTGA 1592
|||||
```



**THIS PAGE BLANK (USPTO)**

QY	19	GAGCAGCGGGGGTGTGTCCTGTGTTGGCCATGGCCGACTACTGTAATTAATGGGGGCAAGT	78
Db	1	GAGCAGACGGCGGGTGTGTCCTGTGTTGGCCATGGCCGACTACTGTAATTAATGGGGGCAAGT	60
QY	79	CCTACGTGGCCGACGACGAGGACTGACAGCAAGCGCTTCAATGTGGGAGACGGCCCTCA	138
Db	61	CCTACGTGGCCGACGAGGACTGACAGCAAGCGCTTCAATGTGGGAGACGGCCCTCA	120
QY	139	CCTACATGACTTTTTCATTTCCCTGGGTACATGACTTCACTGACGACCAAGGTGGACC	198
Db	121	CCTACATGACTTTTTCATTTCCCTGGGTACATGACTTCACTGACGACCAAGGTGGACC	180
QY	199	TGACTTTCGCTTCATCCAAAGAAATTCACTGTTAGACCCCACTGGTTCTCTCCCAATGG	258
Db	181	TGACTTTCGCTTCATCCAAAGAAATTCACTGTTAGACCCCACTGGTTCTCTCCCAATGG	240
QY	259	ACACAGTCCAGAGGCGTGGCATGCGCCATAGCAATGCGCGCTTACAGCGCGTATTGGCTTCA	318

```
|||||
Db 241 ACACAGTCACAGAGCTGGGATGGCCATAGCAATGGCCCTACAGCCGCTATTGGCCCTCA 300
OY 319 TCCACCCAACTGATACACCTGAATCCAGGGCAATGAAGTTGCGAAATGAAATATG 378
Db 301 TCCACCCAACTGATACACCTGAATCCAGGGCAATGAAGTTGCGAAATGAAATATG 360
OY 379 AACAGGATTCATGACAGACCCCTGTGTCTCAGCCCCAGAGATCGCTGGCGGATGTTT 438
Db 361 AACAGGATTCATGACAGACCCCTGTGTCTCAGCCCCAGAGATCGCTGGCGGATGTTT 420
OY 439 TTGAGGCCAAGGCCCGCATGGTTTCTCGCGGTATCCCAATCAGACACAGGCCGGATGG 498
Db 421 TTGAGGCCAAGGCCCGCATGGTTTCTCGCGGTATCCCAATCAGACACAGGCCGGATGG 480
OY 499 GGAGCGGCTTGGTGGGATCATCTCCAGGAGCATTTGTTTTCAAAAGAGAGAAC 558
Db 481 GGAGCGGCTTGGTGGGATCATCTCCAGGAGCATTTGTTTTCAAAAGAGAGAAC 540
OY 559 ATGACTGTTTCTTGGAGAGATGAATGACAAAGAGGAGACTTGTGTAGCCCCCGCA 618
Db 541 ATGACTGTTTCTTGGAGAGATGAATGACAAAGAGGAGACTTGTGTAGCCCCCGCA 600
OY 619 GCATCACACTGAGAGAGGCAATGAATTTCTGACCGCAGCAAGAGGAAAGTTGCCA 678
Db 601 GCATCACACTGAGAGAGGCAATGAATTTCTGACCGCAGCAAGAGGAAAGTTGCCA 660
OY 679 TTGTAAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGGAGACCTGAAGAAGATC 738
Db 661 TTGTAAATGAAGATGATGAGCTTGTGGCCATCATTTGCCGGAGACCTGAAGAAGATC 720
OY 739 GGGACTACCCACTAGCTAGCTCCCAAGATGCCAAGAAACAGCTGCTGTGTGGGAGCCATTG 798
Db 721 GGGACTACCCACTAGCTAGCTCCCAAGATGCCAAGAAACAGCTGCTGTGTGGGAGCCATTG 780
OY 799 GCATCATGAGATGACAAAGTATAGGCTGACTTGTGCCCCAGGCTGCTGTGATGTAG 858
Db 781 GCATCATGAGATGACAAAGTATAGGCTGACTTGTGCCCCAGGCTGCTGTGATGTAG 840
OY 859 TGGTTTTGGACTCTTCCAGGGGAAATTCATCTTCAGATCATATGATCAAGTACATCA 918
Db 841 TGGTTTTGGACTCTTCCAGGGGAAATTCATCTTCAGATCATATGATCAAGTACATCA 900
OY 919 AAGACAAATACCTTAATCTCAAGTCAATTGGAAGCAATGTGCTACTGTGCCCGACCA 978
Db 901 AAGACAAATACCTTAATCTCAAGTCAATTGGAAGCAATGTGCTACTGTGCCCGACCA 960
OY 979 AAGACCTCATTTGATGAGGTGTGATGCCCTCGGGTGGGATGGAAGTGGCTCATCT 1038
Db 961 AAGACCTCATTTGATGAGGTGTGATGCCCTCGGGTGGGATGGAAGTGGCTCATCT 1020
OY 1039 GCATTATCCAGGAAGTGTGGCTGTGGGCGGCCCAAGCAACAGCAGTGTACAAAGTGT 1098
Db 1021 GCATTATCCAGGAAGTGTGGCTGTGGGCGGCCCAAGCAACAGCAGTGTACAAAGTGT 1080
OY 1099 ATGAGTATGACAGCGGCTTGTGTCCGTCAATGCTGATGAGGAATTCAAAATGTGG 1158
Db 1081 CAGAGTATGACAGCGGCTTGTGTCCGTCAATGCTGATGAGGAATTCAAAATGTGG 1140
OY 1159 GTCATATTGCGAAAGCTTGGCCCTTGGGGCTCCACAGTCATGATGGGCTCTCTCTGG 1218
Db 1141 GTCATATTGCGAAAGCTTGGCCCTTGGGGCTCCACAGTCATGATGGGCTCTCTCTGG 1200
OY 1219 CTGGCCACCACTGAGGCCCTGTGATACTTCTTTCCGATGGATCCGGCTAAAGAAAT 1278
Db 1201 CTGGCCACCACTGAGGCCCTGTGATACTTCTTTCCGATGGATCCGGCTAAAGAAAT 1260
OY 1279 ATGCGCGTATGGGCTTCTCGATGCCATGAGCAAGCAACCTCAGCAGCCAGACAGATATT 1338
Db 1261 ATGCGCGTATGGGCTTCTCGATGCCATGAGCAAGCAACCTCAGCAGCCAGACAGATATT 1320
OY 1339 TCAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTGTGTGCTGCAGAGCAAAAG 1398
|||||
```

```
Db 1321 TCAGTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTGTGTGCTGTGCAGGCAAAAG 1380
OY 1399 GGTCAATCCACAAATTTGTGCTTACCTGATTTGCTGSCATCCAACTCATGCCAGACA 1458
Db 1381 GGTCAATCCACAAATTTGTGCTTACCTGATTTGCTGSCATCCAACTCATGCCAGACA 1440
OY 1459 TTGTGTCCAAAGCTTTGACCCCAAGTCCGAGCATGATGCTACTGTGGGAGCTTAAGTTG 1518
Db 1441 TTGTGTCCAAAGCTTTGACCCCAAGTCCGAGCATGATGCTACTGTGGGAGCTTAAGTTG 1500
OY 1519 AGAAGAGAGCTTCTCAGCCAGTGAAGTGGGCTGCATAGCTCCATTGATGAGA 1578
Db 1501 ACAAGAGAGCTTCTCAGCCAGTGAAGTGGGCTGCATAGCTCCATTGATGAGA 1560
OY 1579 AGCGGCTTTTCTGAAAAGGATCCAGCACACCTCTCGGTTTTTTTTTCAATAAAGTTT 1638
Db 1561 AGCGGCTTTTCTGAAAAGGATCCAGCACACCTCTCGGTTTTTTTTTCAATAAAGTTT 1620
OY 1639 AGAAGA 1645
Db 1621 AGAAGA 1627
```

## RESULT 2

```
US-09-822-849A-48
: Sequence 48, Application US/09822849A
: Patent No. US2002045170A1
: GENERAL INFORMATION:
: APPLICANT: Wong, Gordon G.
: APPLICANT: Clark, Hilary
: APPLICANT: Fechtel, Kim
: APPLICANT: Agostino, Michael J.
: APPLICANT: Resnick, Steven H.
: APPLICANT: Resnick, Richard J.
: APPLICANT: Gulakota, Kamalakar
: APPLICANT: Graham, James R.
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
: FILE REFERENCE: GIN 6403
: CURRENT APPLICATION NUMBER: US/09/822, 849A
: CURRENT FILING DATE: 2001-09-04
: PRIOR APPLICATION NUMBER: 60/195,582
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 598
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 48
: LENGTH: 1157
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-822-849A-48
```

```
Query Match 69.0%; Score 1141.6; DB 10; Length 1157;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1147; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 489 GCGCGGATGGGAGCCGCTTGTGGGATCATCTCTCCAGGAGCAATGATTTTCCAAA 548
Db 1 GCGCGGATGGGAGCCGCTTGTGGGATCATCTCTCCAGGAGCAATGATTTTCCAAA 60
OY 549 GAGGAGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGAGACTTGTGTGA 608
Db 61 GAGGAGAACATGACTGTTTCTTGAAGAGATATGACAAAGAGGAGACTTGTGTGA 120
OY 609 GCCCCCCGAGCATCAGCTGAAAGGCAATGAATTTCTGCAGCGCAGCAAGAGGA 668
Db 121 GCCCCTGAGGCACTACAGCTGAAGGAGCAAAATGAATTTCTGCAGCGCAGCAAGAGGA 180
OY 669 AAGTTGCCATTTGAATGAAGATGATGAGCTTGTGGCATCATTTGCCGGAGACGCTG 728
Db 181 AAGTTGCCATTTGAATGAAGATGATGAGCTTGTGGCATCATTTGCCGGAGACGCTG 240
OY 729 AAGGAAGTGGGACTACCCACTAGCCTCCAAAGATGCCAAGAGAGCTGTGTGG 788
|||||
```

Db	241	AAGAAAGTAATGGGACATCCCACTAGACCTCCAAAGATGCCAACAACAAACACTGGTGTGTGGG	300
Oy	789	GCACCCATTGGCACTCATAGAGATGACAAAGTATAGCTTGACTGGTGGCCAGGCTGGT	848
Db	301	GCACCCATTGGCACTCATAGAGATGACAAAGTATAGCTTGACTGGTGGCCAGGCTGGT	360
Oy	849	GTGATGTAGTGGTGGTGGAGCTCTCCAGGCAAAATTCATCTTCCCATCATCAATATGATC	908
Db	361	GTGATGTAGTGGTGGTGGAGCTCTCCAGGCAAAATTCATCTTCCCATCATCAATATGATC	420
Oy	909	AAGTACATCAAAGCAAAATACCTTAATCTTCGAAGTATATGGAGCAATGTGTACTCTCT	968
Db	421	AAGTACATCAAAGCAAAATACCTTAATCTTCGAAGTATATGGAGCAATGTGTACTCTCT	480
Oy	969	GCCAGGCCAAGAACCTCATTTGATGCAAGTGTGTGATAGCTCTCCGGTATGSGCATGGGAAGT	1028
Db	481	GCCAGGCCAAGAACCTCATTTGATGCAAGTGTGTGATAGCTCTCCGGTATGSGCATGGGAAGT	540
Oy	1029	GGCTCCATCTGCATTAATCCAGGAAGTGTGGCTGTGGGCGGCCCAACAACAGCACTG	1088
Db	541	GGCTCCATCTGCATTAATCCAGGAAGTGTGGCTGTGGGCGGCCCAACAACAGCACTG	600
Oy	1089	TACAAGTGTATGAGTATGCAACGGCCCTTGTGTCTCCGCTATTGCTATGTGAGAGAAATC	1148
Db	601	TACAAGTGTATGAGTATGCAACGGCCCTTGTGTCTCCGCTATTGCTATGTGAGAGAAATC	660
Oy	1149	CAAAATGTGGGTCTATTTGGGAAGCCTTGGGCGCCCTGGGGGCCCTCCACATCATATGATGGC	1208
Db	661	CAAAATGTGGGTCTATTTGGGAAGCCTTGGGCGCCCTGGGGGCCCTCCACATCATATGATGGC	720
Oy	1209	TCTCTCTGGCTGGCACACACTGAGGCCCCCTGTGTAATCTTTTCCGATGAGATCCGG	1268
Db	721	TCTCTCTGGCTGGCACACACTGAGGCCCCCTGTGTAATCTTTTCCGATGAGATCCGG	780
Oy	1269	CTTAAGAAATATCGCGGTATGGGCTCTCTCGATGCCATGAGCAACACTCTGACAGCAG	1328
Db	781	CTTAAGAAATATCGCGGTATGGGCTCTCTCGATGCCATGAGCAACACTCTGACAGCAG	840
Oy	1329	AACGATATTTTCACTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTCTGTCTGTG	1388
Db	841	AACGATATTTTCACTGAAGCTGACAAATCAAAAGTGGCCAGGAGTGTCTGTCTGTG	900
Oy	1389	CAGACACAAGGGTCAATCCCAAAATTTGTCCCTTACCTATTGTGGGCTCCCAACATCA	1448
Db	901	CAGACACAAGGGTCAATCCCAAAATTTGTCCCTTACCTATTGTGGGCTCCCAACATCA	960
Oy	1449	TGCCAGACATTTGTGCCAAGACCTTGACCCAAAGTCCGAGCCATATGTACTTGGGAG	1508
Db	961	TGCCAGACATTTGTGCCAAGACCTTGACCCAAAGTCCGAGCCATATGTACTTGGGAG	1020
Oy	1509	CTTAACTTTTGAAGAAGACAGTCTCTAGCCCAAGGTGGAAGTGGGCTCATAGCCTCAT	1568
Db	1021	CTTAACTTTTGAAGAAGACAGTCTCTAGCCCAAGGTGGAAGTGGGCTCATAGCCTCAT	1080
Oy	1569	TTCGATAGAGAAGCGGCTTTTCTCTAAAAAGGATCCAGACACTCTCGCTTTTCTTCA	1628
Db	1081	TTCGATAGAGAAGCGGCTTTTCTCTAAAAAGGATCCAGACACTCTCGCTTTTCTTCA	1140
Oy	1629	ATAAAGCTTTAGAAG 1644	
Db	1141	ATAAAGCTTTAGAAG 1156	
RESULT 3			
US-09-880-107-2119			
Sequence 2119, Application US/09880107			
Patent No. US20020142981A1			
GENERAL INFORMATION:			
: APPLICANT: Horne, Darci T.			
: APPLICANT: Vockley, Joseph G.			
: APPLICANT: Scherf, Uwe			
: APPLICANT: Gene Logic, Inc.			
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			

[illegible]

QY	817	AGTATACCGAGACCTTGCCTGCCAGGCTGGTGGATGATGCTTGTGGACCTTCC	876
Db	1370	AATACCGCTGAGACCTGCTGACCCAGGCGGGGCTGACGCTCATAGCTTCCACTGCTCC	1429
QY	877	AGGGAATTCATCTTCCAGATCAATATGATCAAGTACATCAAGACAATACCTTAATC	936
Db	1430	AAGGAATTCGGTGTATCAATATCGCATCGTGCATTAATCAACAGAAATGACCCACC	1489
QY	937	TCCAACTCATTTGGAGCCAAATGTGTGTCACGTGTGCCAGGCCAAGAACTCATATGATGAG	996
Db	1490	TCCAGGTGATTTGGGGGGAAGCTGTGTACAGCAGCCAGGCCAAGAACTGATGTGCTG	1549
QY	997	GTTGGATGTGCCCTCTCGGGGTGGGCGATGGGAAGTGCTCCATCTGCATTAATCCAGAAATGC	1056
Db	1550	GTTGGAGCGGGGCTGCGGGGTGGGCGATGGGCTCCGCGCTCCATCTGCATCAOCCAGGAATGA	1609
QY	1057	TGGCCTTGGGGGGGGCCCCAAGCAACAGCAGGTATACAGGTATATAGTATGACGGGCT	1116
Db	1610	TGGCGCTTGGTGGCCCCCAGGGCAGCTGCTGTACAAAGGTGGCTGAGTATGGCCGGCTT	1669
QY	1117	TTTGGTGTCCGGGATCATGTCGTATGAGAGGAATCCAAATTTGGGCTATATTGGCAAGGCT	1176
Db	1670	TTGGTGTGCCCATCATATAGCCGATGGCGGCATCCAGACCGTGGGCAACGTGGTCAAGGCC	1729
QY	1177	TGGCGCTTGGGGGCTCCACAGTCATGATGAGCTCTCTCTGGCTGCCACACACTGAGGCC	1236
Db	1730	TGGCGCTTGGAGCCTCCACAGTATGATGATGAGCTCCCTGCTGGCCGCCACTACGAGAGGCC	1789
QY	1237	CTGTGGTAATCTCTTTCCGATGGGATCCGGCTAAAGAAATTTGGCGTATGGGCTTC	1286
Db	1790	CTTGGCGAGTACTCTTCTTCAGACGGGGTGGGCTCAAGAAATGACGGGGGCTGCTAC	1849
QY	1297	TGATGTGCATGAGACAAACACCTCAGCAGGCCAGAAAGATATTTCAAGTCAAGCTCAAAA	1356
Db	1850	TGATATCCATGAGAAAGAGCAGCAGCAGCAAAAGATACTTCAAGCGAGGGGCGATTAAG	1909
QY	1357	TCAAGTGGCCCCAGGAGGTCTGTGTGTGTGTGAGACAAAGGTCATCCCAATTTG	1416
Db	1910	TGAAGATTCGACAGGGGTCTCTGGGCTCCATCCAGGACAAAGATTCATTCGAAGTTG	1969
QY	1417	TCCCTTACCTGATGTGCTGGGATCCAAACACTCATGAGCCAGACATTTGGTGCAGAGCTTGA	1476
Db	1970	TGCCCTACCTCATATGACGAGGATCCAAACAGGGCTGCCAAGATATTCGGGGGCCGACCTGT	2029
QY	1477	CCCAAGTCCGAGCCATGATGATCTCTGGGAGCTTAACTTTGAGAGAGACGCTCTCAG	1536
Db	2030	CTGTCTTCCTGGTCCATCATATGATGACTACAGAGAGCTCAAGTTTGAAGACCGACATGTGCG	2089
QY	1537	CCGAGGTGAGAGGTGGCGTCCATAGGCTCCATTCGATGAGAGCGGCTTTTCTGA	1592
Db	2090	CCCAAGATTGAGGGTGTGTCTCATGTGCTGACTCTTACGAAAGCGGCTGTACTGA	2145
RESULT 4			
US-09-853-918-45			
: Sequence 45: Application us/09853918			
: Patent No. US20020068346A1			
: GENERAL INFORMATION:			
: APPLICANT: Krystek, Stanley R.			
: APPLICANT: Sheriff, Steven			
: APPLICANT: Wiltner, Mark R.			
: APPLICANT: Hollenbaugh, Diane L.			
: APPLICANT: Yan, Ning			
: APPLICANT: Mouravielf, Julie E.			
: APPLICANT: Einspahr, Howard M.			
: APPLICANT: Kish, Kevin			
: TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE			
: FILE REFERENCE: DB24NP			
: CURRENT APPLICATION NUMBER: us/09/853,918			
: PRIOR APPLICATION NUMBER: 60/203,448			
: PRIOR FILING DATE: 2000-05-10			

[illegible]





|||||  
Db 745 GCCCTGGGGCTCCACAGATCATGATGGCTCTCTCCGCGGCCACCACTGAGGCCCT 804  
Oy 1239 GGTGAATCTCTTTCCGATGGATCCGGCTAAAGAAATATCCGGTATGGCTCTC 1298  
|||||  
Db 805 GGTGATCTCTTTCCGATGGATCCGGCTAAAGAAATATCCGGTATGGCTCTC 864  
Oy 1239 GATGCCATGACAAACACCTCAGCAGCCAGAAACAGATATTTTCAGTGAAGCTGACAAATC 1358  
|||||  
Db 865 GATGCCATGACAAACACCTCAGCAGCCAGAAACAGATATTTTCAGTGAAGCTGACAAATC 924  
Oy 1339 AAATGGCCCGAGGAGTGTCTGTCTGCGAGGACAAAGGTCATCCACAAATTTGTC 1418  
|||||  
Db 925 AAAGTGGCCCGAGGAGTGTCTGTCTGCGAGGACAAAGGTCATCCACAAATTTGTC 984  
Oy 1419 CCTTACCTGATGCTGGGATCCACACTCATGCGAGCATTTGGGCCAAGACTTTGACC 1478  
|||||  
Db 985 CCTTACCTGATGCTGGGATCCACACTCATGCGAGCATTTGGGCCAAGACTTTGACC 1044  
Oy 1479 CAAGTCCGAGCCATGATGCTCTGGGAGCTTAAAGTTTGAGAGAGAAAGCTCCAGCC 1538  
|||||  
Db 1045 CAAGTCCGAGCCATGATGCTCTGGGAGCTTAAAGTTTGAGAGAGAAAGCTCCAGCC 1104  
Oy 1539 CAGGTGGAAGTGGGCTCATAGCTCCATTGCTATGAGAGAGCGCTTTCTGA 1592  
|||||  
Db 1105 CAGGTGGAAGTGGGCTCATAGCTCCATTGCTATGAGAGAGCGCTTTCTGA 1158

## RESULT 7

US-09-853-918-40

; Sequence 40, Application US/09853918

; Patent No. US2002068346A1

; GENERAL INFORMATION:

; APPLICANT: Krystek, Stanley R.

; APPLICANT: Sherif, Steven

; APPLICANT: Wilmer, Mark R.

; APPLICANT: Hollenbaugh, Diane L.

; APPLICANT: Yan, Ning

; APPLICANT: Mouravieff, Julie E.

; APPLICANT: Einspahr, Howard M.

; APPLICANT: Kish, Kevin

; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE

; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: DB24NP

; CURRENT APPLICATION NUMBER: US/09/853,918

; CURRENT FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: 60/203,448

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 40

; LENGTH: 1155

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-853-918-40

Query Match 48.9%; Score 809.6; DB 10; Length 1155;  
Best Local Similarity 99.5%; Pred. No. 2.8e-243;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 777 CTGCTGTGGGGGACCCATTGGCACTCATGAGAGTGAACAATATAGGCTGGCACTTCTC 836  
|||||  
Db 340 CTGCTGTGGGGGACCCATTGGCACTCATGAGAGTGAACAATATAGGCTGGCACTTCTC 399  
Oy 837 GCCCAGCGGTGGTGAATAGTGGTTTGGACTCTTCCAGGAAATTCATCTCCAG 896  
|||||  
Db 400 GCCCAGCGGTGGTGAATAGTGGTTTGGACTCTTCCAGGAAATTCATCTCCAG 459  
Oy 897 ATCAATATGATCAAGTACATCAAGACAATACCTTAATCTCCAGTCAATTTGGAGCAAT 956  
|||||  
Db 460 ATCAATATGATCAAGTACATCAAGACAATACCTTAATCTCCAGTCAATTTGGAGCAAT 519  
Oy 957 GTGTGCTACTGCTCCAGGCCAAGAACTCATGATGACAGTGTGGATCCCTGCGGCTG 1016

|||||  
Db 520 GTGTGCTACTGCTCCAGGCCAAGAACTCATGATGACAGTGTGGATGCCCTGCGGCTG 579  
Oy 1017 GGCATGGAATGGCTCCATCTGCAATTTCCAGAAAGTCTGAGCTTGGGCGGCCCA 1076  
|||||  
Db 580 GGCATGGAATGGCTCCATCTGCAATTTCCAGAAAGTCTGAGCTTGGGCGGCCCA 639  
Oy 1077 GCACAGCAGTGTACAAAGTGTATGATGATGACAGCGCTTTGGTTTCCGCTATTTCT 1136  
|||||  
Db 640 GCACAGCAGTGTACAAAGTGTATGATGATGACAGCGCTTTGGTTTCCGCTATTTCT 699  
Oy 1137 GATGAGAAATCCAAATGTGGGTCTATTTGGAAAGCTTGGGCTTGGGCGGCCCA 1196  
|||||  
Db 700 GATGAGAAATCCAAATGTGGGTCTATTTGGAAAGCTTGGGCTTGGGCGGCCCA 759  
Oy 1197 GTCATGATGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1256  
|||||  
Db 760 GTCATGATGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 819  
Oy 1257 GATGGATCCGCTTAAAGAAATATGCGGCTATGGCTTCTCTGATGCTGATGCTGATGCT 1316  
|||||  
Db 820 GATGGATCCGCTTAAAGAAATATGCGGCTATGGCTTCTCTGATGCTGATGCTGATGCT 879  
Oy 1317 CTCACAGCCCAAGACATATTTTCAAGTGAAGTGAACAAATCAAGTGGCCAGGAGTG 1376  
|||||  
Db 880 CTCACAGCCCAAGACATATTTTCAAGTGAAGTGAACAAATCAAGTGGCCAGGAGTG 939  
Oy 1377 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1436  
|||||  
Db 940 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999  
Oy 1437 ATCCAACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1496  
|||||  
Db 1000 ATCCAACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1059  
Oy 1497 TACTCTGGGAGCTTAAAGTGTGAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1556  
|||||  
Db 1060 TACTCTGGGAGCTTAAAGTGTGAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119  
Oy 1557 CATAGCTTCATTCGTATGAGAAAGCGGCTTTCTGA 1592  
|||||  
Db 1120 CATAGCTTCATTCGTATGAGAAAGCGGCTTTCTGA 1155

## RESULT 8

US-09-853-918-41

; Sequence 41, Application US/09853918

; Patent No. US2002068346A1

; GENERAL INFORMATION:

; APPLICANT: Krystek, Stanley R.

; APPLICANT: Sherif, Steven

; APPLICANT: Wilmer, Mark R.

; APPLICANT: Hollenbaugh, Diane L.

; APPLICANT: Yan, Ning

; APPLICANT: Mouravieff, Julie E.

; APPLICANT: Einspahr, Howard M.

; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE

; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: DB24NP

; CURRENT APPLICATION NUMBER: US/09/853,918

; CURRENT FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: 60/203,448

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 41

; LENGTH: 1155

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-853-918-41

Query Match 48.9%; Score 809.6; DB 10; Length 1155;

Best Local Similarity 99.5%; Pred. No. 2,8e-243;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 777 CTGCTGTGGGGGACCCATTTGGCACTCATGAGATGACAGATATGCTGACCTTGCTC 836
Db 340 CTGCTGTGGGGGACCCATTTGGCACTCATGAGATGACAGATATGCTGACCTTGCTC 399
QY 837 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCAG 896
Db 400 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCAG 459
QY 897 ATCAATATGATCAAGTACATCAAGCAAGCAATACCTTAATCTCCAACTGATTTGAGGCAT 956
Db 460 ATCAATATGATCAAGTACATCAAGCAAGCAATACCTTAATCTCCAACTGATTTGAGGCAT 519
QY 957 GTGGTCACTGCTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGACCCCTGGGGGTG 1016
Db 520 GTGGTCACTGCTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGACCCCTGGGGGTG 579
QY 1017 GCCATGGGAAGTGGCTTCATCTTCATTTATCCAGGAAGTCTGGCCCTGCGCGGCCCAA 1076
Db 580 GCCATGGGAAGTGGCTTCATCTTCATTTATCCAGGAAGTCTGGCCCTGCGCGGCCCAA 639
QY 1077 GCACAGCAGTGTACAGGTGTATGATGTGACAGGCGCTTGGTGTGCGGTCAATTGCT 1136
Db 640 GCACAGCAGTGTACAGGTGTATGATGTGACAGGCGCTTGGTGTGCGGTCAATTGCT 699
QY 1137 GATGGAGGAATCCAAATGTGGTTCATTTGCGAAAGCCTTGGCCCTTGGGGCTTCACA 1196
Db 700 GATGGAGGAATCCAAATGTGGTTCATTTGCGAAAGCCTTGGCCCTTGGGGCTTCACA 759
QY 1197 GTCATGATGGCTCTCTCTGCTGCCACACCTGAGGCGCTTGGTGAATCTTCTTTTC 1256
Db 760 GTCATGATGGCTCTCTCTGCTGCCACACCTGAGGCGCTTGGTGAATCTTCTTTTC 819
QY 1257 GATGGATCCGGCTAAAGAAATATCCGGTATGGGTCTCTGATCCCATGCAACAGCAC 1316
Db 820 GATGGATCCGGCTAAAGAAATATCCGGTATGGGTCTCTGATCCCATGCAACAGCAC 879
QY 1317 CTCAGCAGCAGAACAGATATTTCACTGAAGCTGCACAAATCAAACTGACCCAGGAGTG 1376
Db 880 CTCAGCAGCAGAACAGATATTTCACTGAAGCTGCACAAATCAAACTGACCCAGGAGTG 939
QY 1377 TCTGGTGTGTGAGAGCAAAAGGTCATCCAAATTTGTCCTTACCTGATTTGCTGGC 1436
Db 940 TCTGGTGTGTGAGAGCAAAAGGTCATCCAAATTTGTCCTTACCTGATTTGCTGGC 999
QY 1437 ATCCAACTCATTCGCAGACATTTGGTGCAGAGCTTGACCCAACTG 1496
Db 1000 ATCCAACTCATTCGCAGACATTTGGTGCAGAGCTTGACCCAACTG 1059
QY 1497 TACTCTGGGAGCTTAAGTTTGAGAAGAGAGCTTCTCCAGCCCAAGTG 1556
Db 1060 TACTCTGGGAGCTTAAGTTTGAGAAGAGAGCTTCTCCAGCCCAAGTG 1119
QY 1557 CATAGCCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1592
Db 1120 CATAGCCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1155
```

## RESULT 9

```
US-09-853-918-42
; Sequence 42, Application US/09853918
; Patent No. US20020068346A1
; GENERAL INFORMATION:
; APPLICANT: Krystek, Stanley R.
; APPLICANT: Sheriff, Steven
; APPLICANT: Witmer, Mark R.
; APPLICANT: Hollenbaugh, Diane L.
; APPLICANT: Yan, Ning
; APPLICANT: Mouravieff, Julie E.
; APPLICANT: Einspahr, Howard M.
; APPLICANT: Kish, Kevin
```

```
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: DB24NP
; CURRENT APPLICATION NUMBER: US/09/853,918
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,448
; PRIORITY FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-918-42
```

Query Match 48.9%; Score 809.6; DB 10; Length 1155;

Best Local Similarity 99.5%; Pred. No. 2,8e-243;  
Matches 812; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 777 CTGCTGTGGGGGACCCATTTGGCACTCATGAGATGACAGATATGCTGACCTTGCTC 836
Db 340 CTGCTGTGGGGGACCCATTTGGCACTCATGAGATGACAGATATGCTGACCTTGCTC 399
QY 837 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCAG 896
Db 400 GCCCAGGCTGGTGTGATGTAGTGTGTTGGACTCTTCCAGGGAATTCATCTTCAG 459
QY 897 ATCAATATGATCAAGTACATCAAGCAAGCAATACCTTAATCTCCAACTGATTTGAGGCAT 956
Db 460 ATCAATATGATCAAGTACATCAAGCAAGCAATACCTTAATCTCCAACTGATTTGAGGCAT 519
QY 957 GTGGTCACTGCTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGACCCCTGGGGGTG 1016
Db 520 GTGGTCACTGCTGCCAGGCCAAGAACCTCATTTGATGACAGTGTGATGACCCCTGGGGGTG 579
QY 1017 GCCATGGGAAGTGGCTTCATCTTCATTTATCCAGGAAGTCTGGCCCTGCGCGGCCCAA 1076
Db 580 GCCATGGGAAGTGGCTTCATCTTCATTTATCCAGGAAGTCTGGCCCTGCGCGGCCCAA 639
QY 1077 GCACAGCAGTGTACAGGTGTATGATGTGACAGGCGCTTGGTGTGCGGTCAATTGCT 1136
Db 640 GCACAGCAGTGTACAGGTGTATGATGTGACAGGCGCTTGGTGTGCGGTCAATTGCT 699
QY 1137 GATGGAGGAATCCAAATGTGGTTCATTTGCGAAAGCCTTGGCCCTTGGGGCTTCACA 1196
Db 700 GATGGAGGAATCCAAATGTGGTTCATTTGCGAAAGCCTTGGCCCTTGGGGCTTCACA 759
QY 1197 GTCATGATGGCTCTCTCTGCTGCCACACCTGAGGCGCTTGGTGAATCTTCTTTTC 1256
Db 760 GTCATGATGGCTCTCTCTGCTGCCACACCTGAGGCGCTTGGTGAATCTTCTTTTC 819
QY 1257 GATGGATCCGGCTAAAGAAATATCCGGTATGGGTCTCTGATCCCATGCAACAGCAC 1316
Db 820 GATGGATCCGGCTAAAGAAATATCCGGTATGGGTCTCTGATCCCATGCAACAGCAC 879
QY 1317 CTCAGCAGCAGAACAGATATTTCACTGAAGCTGCACAAATCAAACTGACCCAGGAGTG 1376
Db 880 CTCAGCAGCAGAACAGATATTTCACTGAAGCTGCACAAATCAAACTGACCCAGGAGTG 939
QY 1377 TCTGGTGTGTGAGAGCAAAAGGTCATCCAAATTTGTCCTTACCTGATTTGCTGGC 1436
Db 940 TCTGGTGTGTGAGAGCAAAAGGTCATCCAAATTTGTCCTTACCTGATTTGCTGGC 999
QY 1437 ATCCAACTCATTCGCAGACATTTGGTGCAGAGCTTGACCCAACTG 1496
Db 1000 ATCCAACTCATTCGCAGACATTTGGTGCAGAGCTTGACCCAACTG 1059
QY 1497 TACTCTGGGAGCTTAAGTTTGAGAAGAGAGCTTCTCCAGCCCAAGTG 1556
Db 1060 TACTCTGGGAGCTTAAGTTTGAGAAGAGAGCTTCTCCAGCCCAAGTG 1119
QY 1557 CATAGCCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1592
Db 1120 CATAGCCTTCATTCGTATGAGAAGCGGCTTTTCTGA 1155
```





```
RESULT 13
US-09-853-918-44
: Sequence 44, Application US/09853918
: Patent No. US20020068346A1
: GENERAL INFORMATION:
: APPLICANT: Kryslek, Stanley R.
: APPLICANT: Sheriff, Steven
: APPLICANT: Witmer, Mark R.
: APPLICANT: Hollenbaugh, Diane L.
: APPLICANT: Yan, Ning
: APPLICANT: Moutavieff, Julie E.
: APPLICANT: Einspahr, Howard M.
: APPLICANT: Kish, Kevin
: TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE
: TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: DB24NP
: CURRENT APPLICATION NUMBER: US/09/853,918
: CURRENT FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: 60/203,448
: PRIOR FILING DATE: 2000-05-10
: NUMBER OF SEQ ID NOS: 65
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 44
: LENGTH: 1155
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-853-918-44
```

```
Query Match      31.1%; Score 513.6; DB 10; Length 1155;
Best Local Similarity 76.5%; Pred. No. 1.7e-150;
Matches 630; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
```

```
QY 769 AGAAGACGTGCTGGTGTGGGCGACCATTTGGCATGTAGAGTCAAGTATTAAGCTGG 828
DB 332 AAAAAACCTGCTGCTGTGTGGGACAGCTGTGGGACCCGTGAGGTGCAATATCCGTCTGG 391
QY 829 ACTTGCTCCGCCAGAGCTGTGTGGATGTAGTGTGTTGGACTCTTCCAGGAAATTCCA 888
DB 392 ACCTGCTACCCAGGCGGGGCTGTAGCTAGTCTTGGACTGTGTCGAAGGAATTCCG 451
QY 889 TCTTCCAGATCATATATGATCAAGTACATCAAGACAATAATCCCTAATCTCCAAGTATTG 948
DB 452 TGTATCAAAATCCCATGTGTCATTTACATCAAAACAGAAATGCCCCCTCCAGGTATTG 511
QY 949 GAGGCAATGTGGTCACTGTGTCGCCAGGCCAAGAACTCATTTGATGAGGTGGATGCC 1008
DB 512 GGGGGAACGTGTGTGACAGACGCCAGGCCAAGAACTGATTGATGCTGTGTGGACGGC 571
QY 1009 TCGGGTGGGTCATGTGAAGTGGCTCCATCTGCATTTATCCAGAAAGTGTGGCTGTGGGC 1068
DB 572 TCGGCTGTGGGTCATGTGGCTCGGGCTCATCTGCATCCAGGAAGTATGGCTGTGGTC 631
QY 1069 GGGCCCAACAGACAGAGTGTACAAAGTGTATGAGTATGACAGGGCCCTTGTGTCCG 1128
DB 632 GGCCCCAGGGCACTGTGTGTACAAAGTGTGAGTATGCCCCCTTGTGTGCCA 691
QY 1129 TCATTGCTATGAGGAGTAATCAAAATGTGGTTCATTTGCGAAAGCTTGGCCCTTGGGG 1188
DB 692 TCATTAGCCATGGCGGTCATCCAGACCGCTGGGACACGTGTCAAGGGCTGGCCCTTGGAG 751
QY 1189 CCTCCACAGTATGATGGGCTCTCTCCCTGTCGCCACACTAGGCCCTCGTGAATATCT 1248
DB 752 CCTCCACAGTATGATGGGCTCTCTCCCTGTCGCCACACTAGGCCCTCGTGGGAATACT 811
QY 1249 TCTTTTCGATGGGATCGGGCTTAAGAAATATTCGGGGTATGGGTTCTCTCGATGCCATGG 1308
DB 812 TCTTTCACAGCGGGGTGCGCTCAAGAAATACCGGGGTCATGGGCTCACTGATGCATGG 871
QY 1309 ACAAGCACTCAGCAGCCAGAAACAGATATTTGATGATGACCTGACAAATCAAGTGGCCC 1368
DB 872 AGAAGAGCAGACAGCAGCAAGAAACGATACCTTACGCGAGGGGGATTAAGTGAAGATCGCGC 931
```

```
QY 1369 AGGAGTGTCTGTGCTGTGACAGACAAAGGTCATTCACAATTTGTCTCTACTGA 1428
DB 932 AGGCTCTCGGGGCTTCATCCAGACAAAGATTCATTCAGAAGTTCGTGGCTTACTCA 991
QY 1429 TTGCTGGCATCCACACATCATGCGCAGGACATTTGTGCCAAGAGCTTGACCCAAAGTCCGAG 1488
DB 992 TAGCAGGCATCCACACAGGCTCTCCAGGATATCGGGGCCGCGCTGTCTGTCTTCGGT 1051
QY 1489 CCATGATGTACTCTGGGAGCTTTAAGTTGAGAGAAGAGACGTCTCAGCCGACGAGGAAG 1548
DB 1052 CCATGATGTACTCAGAGAGAGCTTCAGATTGAGAGAAGGACCATGTGCGGCCAGATTGAG 1111
QY 1549 GTGGCGTCCATAGCCTTCATTCGTATGACAGCGGCTTTCTGA 1592
DB 1112 GTGGTCTCATGCGCTGCACCTTACGAAAGCGGCTGTACTGA 1155
```

```
RESULT 14
US-09-879-536-844
: Sequence 844, Application US/09879536
: Patent No. US20020144298A1
: GENERAL INFORMATION:
: APPLICANT: Endege, Wilson O.
: APPLICANT: Steimann, Kathleen E.
: APPLICANT: Astle, Jon H.
: APPLICANT: Burgess, Christopher C.
: APPLICANT: Bushnell, Steven E.
: APPLICANT: Carroll III, Eddie
: APPLICANT: Catino, Theodore J.
: APPLICANT: Dertl, Adnan
: APPLICANT: Ford, Donna M.
: APPLICANT: Lewis, Marcia E.
: APPLICANT: Monahan, John E.
: APPLICANT: Schlegel, Robert
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS
: FILE REFERENCE: CCD-257 (US)
: CURRENT APPLICATION NUMBER: US/09/879,536
: CURRENT FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/088,801
: PRIOR FILING DATE: 1998-06-10
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 844
: LENGTH: 675
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(675)
: OTHER INFORMATION: n = A,T,C or G
US-09-879-536-844
```

```
Query Match      25.4%; Score 420.6; DB 10; Length 875;
Best Local Similarity 95.8%; Pred. No. 1.7e-121;
Matches 432; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 331 GTACACCTGAATTTCCAGGCGCAATGAAGTTCGGAAGTGAAGAAATATGAACAGGGATTCA 390
DB 2 GTACACCTGAATTTCCAGGCGCAATGAAGTTCGGAAGTGAAGAAATATGAACAGGGATTCA 61
QY 391 TCACAGACCTTGTGTGCTCCAGCCCAAGAGTGTGCGGCGATGTTTGTGAGGCCAAG 450
DB 62 TCACAGACCTTGTGTGCTCCAGCCCAAGAGTGTGCGGCGATGTTTGTGAGGCCAAG 121
QY 451 CCGGCGATGTTTGTGCGGCTATCCCAATCAGACACAGAGCCGATGGGAGCCGCTTGG 510
DB 122 CCGGCGATGTTTGTGCGGCTATCCCAATCAGACACAGAGCCGATGGGAGCCGCTTGG 181
QY 511 TGGGATCATCTTCCTCCAGGAGCATTTGATTTTCTCAAGAGAGAGAACATGACTGTTCT 570
DB 182 TGGGATCATCTTCCTCCAGGAGCATTTGATTTTCTCAAGAGAGAGAACATGACTGTTCT 241
```



**THIS PAGE BLANK (USPTO)**